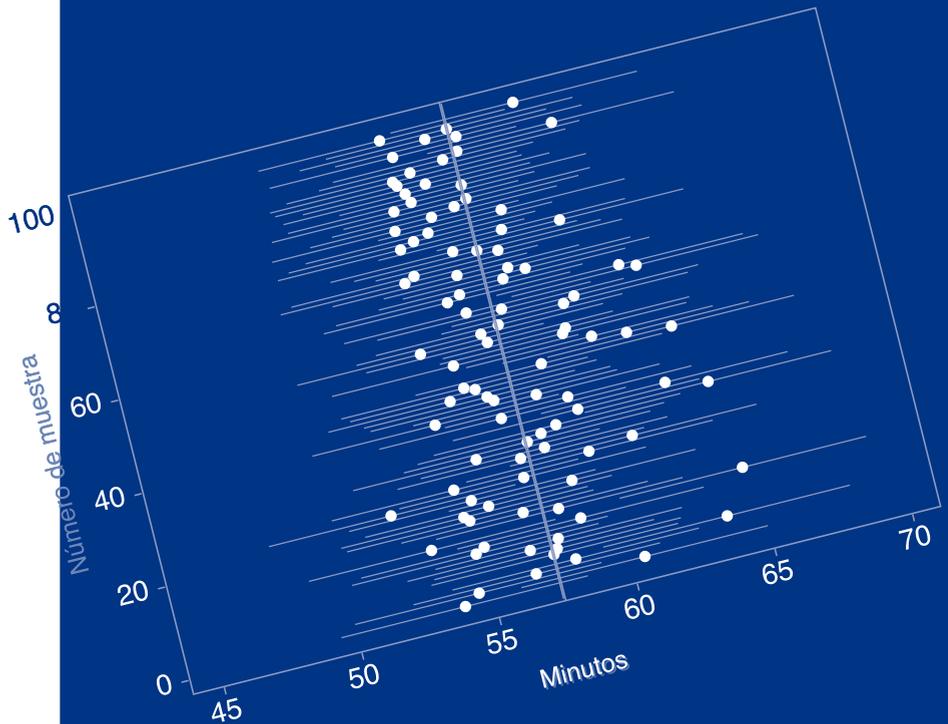


METODOLOGÍA DE INVESTIGACIÓN Y ESCRITURA CIENTÍFICA EN CLÍNICA

Editor:
**Rafael Burgos
Rodríguez**



METODOLOGÍA DE INVESTIGACIÓN Y ESCRITURA CIENTÍFICA EN CLÍNICA

EDITOR:

Rafael Burgos Rodríguez

SERVICIO DE UROLOGÍA
HOSPITAL REGIONAL DE MÁLAGA
SERVICIO ANDALUZ DE SALUD

ESCUELA ANDALUZA
DE SALUD PÚBLICA
GRANADA

1998

Catalogación por la Biblioteca de la EASP

METODOLOGÍA de investigación y escritura científica en clínica / Editor
Rafael Burgos Rodríguez. -3ªed. - Granada: Escuela Andaluza de Salud
Pública, 1998. -(Monografías de la EASP; 24)

1. Metodología de investigación. 2. Publicaciones. I. Burgos Rodríguez,
Rafael, ed.

NLM Classification Q 180
CDU (378) 001

Edita: ESCUELA ANDALUZA DE SALUD PÚBLICA
Campus Universitario de Cartuja. Apdo. de Correos 2070
18080 Granada España

ISBN: 84-87385-45-1



Depósito Legal: Gr-1031/96

Fotocomposición y Gráficos: Miguel Salvatierra

Imprime: Copicentro Granada S.L

Todos los derechos reservados. Ninguna parte de esta publicación puede ser reproducida ni transmitida en ninguna forma ni por ningún medio de carácter mecánico ni electrónico, incluidos fotocopia y grabación, ni tampoco mediante sistemas de almacenamiento y recuperación de información, a menos que se cuente con la autorización por escrito de la Escuela Andaluza de Salud Pública.

Las publicaciones de la Escuela Andaluza de Salud Pública están acogidas a la protección prevista por las disposiciones del Protocolo 2 de la Convención Universal de Derechos de Autor.

Las denominaciones empleadas en esta publicación y la forma en que aparecen presentados los datos que contiene no implican, de parte de la Escuela Andaluza de Salud Pública, juicio alguno sobre la condición jurídica de los países, territorios, ciudades o zonas citados o de sus autoridades, ni respecto a la delimitación de sus fronteras.

La mención de determinadas sociedades mercantiles o del nombre comercial de ciertos productos no implica que la Escuela Andaluza de Salud Pública los apruebe o recomiende con preferencia a otros análogos.

De las opiniones expresadas en la presente publicación responden únicamente los autores.

Los autores:

MANUEL ALEJANDRO BOBENRIETH ASTETE
Doctor en Medicina. Universidad de Chile
Master en Salud Pública. Universidad de Chile
Master en Administración Hospitalaria. Universidad de Minnesota
Coordinador Unidad Técnica de Publicaciones. EASP
Profesor EASP

RAFAEL BURGOS RODRIGUEZ
Médico. Universidad de Granada
Master en Salud Pública y Administración Sanitaria. EASP
Jefe Servicio Urología. Hospital Regional de Málaga
Profesor Asociado. EASP

ADRIANO CALZAS URRUTIA
Psicólogo. Universidad de Granada
Profesor EASP

JOSÉ ANTONIO CHICHARRO MOLERO
Médico. Universidad de Málaga
Responsable Unidad Prostática
Servicio Urología. Hospital Regional de Málaga

RICARDO OCAÑA RIOLA
Matemático. Universidad de Granada
Técnico de Proyectos. EASP

EMILIO PEREA-MILLA LOPEZ
Doctor en Medicina. Universidad de Sevilla
Profesor EASP

EMILIO SANCHEZ-CANTALEJO RAMIREZ
Matemático y Médico. Universidad de Granada
Profesor EASP

ÍNDICE

PRÓLOGO EDICION 1994.....	7
PRÓLOGO EDICION 1996.....	11
AGRADECIMIENTOS PRIMERA EDICIÓN	15
INTRODUCCIÓN A LA EDICIÓN 1996	17
INTRODUCCIÓN A LA PRIMERA EDICIÓN.....	19

PARTE I. DISEÑOS DE INVESTIGACIÓN EPIDEMIOLÓGICA Y SUS APLICACIONES EN CLÍNICA

A. TIPOS DE ESTUDIO	31
1. Conceptos generales.....	31
2. Secuencia del proceso	32
3. Tipos de estudios: la observación y la experimentación	35
4. Estudios experimentales.....	36
5. Estudios no experimentales.....	37
B. ERRORES EN LOS DISEÑOS Y FORMAS DE CONTROL	43
1. Error aleatorio	43
2. Error sistemático o sesgo.....	44
3. Control de los errores	47
C. MEDIDAS DE FRECUENCIA.....	49
1. Introducción.....	49
2. Prevalencia.....	51
3. Incidencia.....	52
4. Indicadores e índices.....	61

D. ESTUDIOS DESCRIPTIVOS.....	63
1. Introducción.....	63
2. Informes de casos y series de casos	64
3. Estudios ecológicos	65
4. Estudios transversales (prevalencia)	69
5. Análisis	72
6. Ventajas.....	74
E. ESTUDIOS DE SEGUIMIENTO	77
1. Introducción.....	77
2. Definición de las cohortes	79
3. Identificación de las cohortes	82
4. Formación de las cohortes	82
5. Determinación de la exposición.....	83
6. Sesgos	83
7. Análisis	85
F. ESTUDIOS DE CASOS Y CONTROLES	95
1. Introducción.....	95
2. Selección de casos.....	96
3. Selección de controles.....	97
4. Grupo control: Población en estudio y base de estudio	101
5. Recogida de la información	102
6. Sesgos	103
7. Análisis de la información	105
G. ESTUDIOS CUASIEXPERIMENTALES Y ESTUDIOS ALEATORIOS CONTROLADOS	111
1. Introducción	111
2. Estudios cuasiexperimentales.....	112
3. Estudios aleatorios controlados o ensayos clínicos.....	116
REFERENCIAS.....	125

PARTE II. ANÁLISIS DE DATOS EN INVESTIGACIÓN CLÍNICA Y EPIDEMIOLOGÍA

A. ANÁLISIS DESCRIPTIVO.....	131
1. Introducción.....	131
2. Medición de signos y síntomas	132
3. Presentación de datos.....	133
4. Resumen numérico de una serie estadística.....	147
5. Otra forma de resumir una serie de datos.....	154
B. TESTS DIAGNÓSTICOS	159
1. Introducción.....	159
2. Probabilidad.....	161
3. Probabilidad condicionada y teorema de Bayes.....	161
4. Influencia de sensibilidad, especificidad y prevalencia en valores predictivos	168
5. Tests diagnósticos y variables continuas	171
6. Aplicaciones de varios tests diagnósticos	174
C. INFERENCIA ESTADÍSTICA. EL ERROR ALEATORIO	177
1. Introducción.....	177
2. De lo particular a lo general.....	177
3. Valor de la información de nuestra experiencia.....	180
4. Control del error aleatorio: Intervalo de confianza.....	183
5. Control del error aleatorio: Contrastes de hipótesis.....	186
6. Tamaño de la muestra	193
D. TESTS PARA COMPARAR DOS GRUPOS.....	197
1. Introducción.....	197
2. Test de la t de Student	198
3. Regresión lineal simple.....	202
4. Correlación	208
5. Regresión logística.....	213
6. Análisis de supervivencia.....	218

E. ANÁLISIS MULTIVARIANTE	229
1. Confusión e interacción	229
2. Necesidad del análisis multivariante	233
3. Regresión lineal múltiple	236
4. Regresión logística múltiple.....	242
5. Análisis de supervivencia: el modelo de Cox.....	247
6. De la asociación a la causalidad.....	253
F. ANÁLISIS DE LA DECISIÓN CLÍNICA.....	259
1. Introducción. Variabilidad.....	259
2. Árbol de decisión	260
3. Probabilidad umbral	269
REFERENCIAS	271

PARTE III. EL SOPORTE INFORMÁTICO EN LA INVESTIGACIÓN

A. INTRODUCCIÓN	277
B. INFORMACIÓN Y DATOS: PREPARACIÓN DE LOS DATOS	279
1. Introducción.....	279
2. Determinación de la unidad de registro	281
3. Diseño de la estructura de registro.....	284
4. Elección del formato de datos	290
C. SOFTWARE DISPONIBLE	295
1. Procesadores o editores de texto	295
2. Hojas de cálculo	297
3. Gestores de bases de datos	298
4. Paquetes estadísticos	299
REFERENCIAS	301

PARTE IV. ESCRITURA Y LECTURA CRÍTICA DE ARTÍCULOS CIENTÍFICOS

A. FORMAS DE COMUNICACIÓN CIENTÍFICA.....	305
1. Tipos de publicaciones científicas	305
2. Estructura de las revistas biomédicas	307
B. LAS ETAPAS DEL PROCESO DE INVESTIGACIÓN Y LA ESTRUCTURA DEL ARTÍCULO CIENTÍFICO ORIGINAL	311
1. Concepto de investigación	311
2. Concepto de científico	313
3. Características del conocimiento científico	314
4. Principios, procesos y métodos.....	317
5. Las etapas del proceso de investigación en salud.....	318
6. Las etapas de la investigación y los elementos del artículo original	323
C. PARTES DEL ARTÍCULO CIENTÍFICO ORIGINAL.....	325
1. Título	326
2. Autores.....	330
3. Resumen	333
4. Introducción.....	336
5. Materiales y métodos	341
6. Resultados.....	345
7. Tablas.....	350
8. Gráficos	370
9. Notas a pie de página.....	405
10. Discusión	408
11. Conclusión	412
12. Agradecimientos	413
13. Referencias	415
14. Apéndices	424
D. PARTES DEL ARTÍCULO DE REVISIÓN	427
1. Concepto y estándares.....	427
2. Pautas para evaluar la calidad de revisión	428
3. Estructura.....	429

E. ESTILO EN LA COMUNICACIÓN CIENTÍFICA	431
1. Redacción, gramática y estilo	431
2. Diferencias entre el estilo científico y el estilo literario.....	433
3. Cualidades del buen estilo	435
4. Papel de la lógica y de la gramática en la redacción correcta.....	437
5. Cincuenta y un consejos útiles para evitar vicios de mal estilo.....	438
F. LECTURA CRÍTICA DE ARTÍCULOS CIENTÍFICOS.....	445
1. El clínico como juez crítico de artículos.....	445
2. Listas de revisión	446
3. Sinopsis de seis pautas de evaluación.....	447
APÉNDICE I: Ejercicio sobre redacción de títulos	453
APÉNDICE II: El proceso de la revisión bibliográfica. Resumen esquemático.....	457
APÉNDICE III: Pautas para evaluación de artículos originales.....	463
Elementos para el análisis crítico de una investigación.....	467
Evaluación de comunicados de investigación	469
Evaluación crítica de la investigación publicada	477
Lista de comprobación para evaluar la investigación publicada	479
Lista de evaluación en la investigación.....	481
Lista de evaluación de artículos médicos generales.....	485
Lista para la evaluación de informes de ensayos clínicos.....	487
APÉNDICE IV: Requisitos uniformes para preparar los manuscritos enviados a revistas biomédicas	489
REFERENCIAS	507
ÍNDICE DE MATERIAS.....	511
ÍNDICE DE AUTORES	521

PRÓLOGO A LA EDICION 1994

Nuestro pensador Miguel de Unamuno, en su libro "El Sentimiento Trágico de la Vida", decía: "La palabra es el cuerpo social del pensamiento". Dificilmente puede expresarse tan concisamente una idea concreta. Mediante la palabra, escrita o hablada, somos capaces los humanos de transmitir hechos trascendentes o, por el contrario, sucesos efímeros y sin importancia.

Bachelard decía "Todo conocimiento es la respuesta a una pregunta", indicando que la curiosidad es uno de los rasgos que comparten hombre y animales. En un sentido amplio, investigar sería tratar de hallar respuestas que satisfagan las preguntas que se plantea el hombre, posiblemente causa de nuestro desconcierto e incertidumbre. En el campo de la ciencia siempre hay problemas de difícil solución, puesto que el conocimiento de lo real es una luz que proyecta inexcusablemente sombras.

Un país tan necesitado como el nuestro de renovaciones debe acoger con entusiasmo todo aquello que signifique progreso e implementación. El texto que ahora prologamos es buena muestra de la respuesta que un grupo de profesionales se plantean cuando tratan de investigar y transmitir el resultado de los conocimientos adquiridos. En una época en la que los usos y costumbres sociales, debido a las nuevas tecnologías, pueden impulsar una evolución de consecuencias imprevisibles, el recuerdo de lo que es una metodología rigurosa y cómo debe transmitirse la misma es un ejemplo a seguir por cualquier intelectual que pretenda ser riguroso en sus exposiciones.

El lenguaje, natural o artificial, conceptual o no conceptual, no sólo es instrumento de comunicación sino que también es, ocasionalmente, instrumento del pensamiento. Quiere ello decir que la escritura científica aporta lo que algunos filósofos opinan sobre las ciencias, es decir, que no son más que un lenguaje artificial. Si no hay una perfecta realización, expresión y transmisión, en el lenguaje escrito ocurrirá desgraciadamente como con aquellos maravillosos murales, que están situados en habitaciones oscuras y de los que no se puede disfrutar de su visión.

Según Bunge, el Planteamiento Científico está constituido por el método científico y el objetivo de la ciencia. Un método científico es un procedimiento para tratar un conjunto de problemas. Cada clase de problemas requiere un conjunto de métodos o técnicas especiales. Los métodos especiales, en la ciencia, pueden ser relevantes para algún estudio particular de la investigación; sin embargo, el método general de la ciencia es un procedimiento que se aplica al ciclo entero de la investigación, en el marco de cada problema de conocimiento. Los pel-

daños principales en el camino de la investigación científica, es decir, de los pasos principales de la aplicación del método científico, puede deducirse una serie ordenada de operaciones:

- 1.- Enunciar preguntas bien formuladas y verosímelmente fecundas.
- 2.- Arbitrar conjeturas fundadas y contrastables con la experiencia para contestar a las preguntas.
- 3.- Derivar consecuencias lógicas de las conjeturas.
- 4.- Arbitrar técnicas para someter las conjeturas a contrastación.
- 5.- Someter a su vez a contrastación esas técnicas, para comprobar su relevancia y la fe que merecen.
- 6.- Llevar a cabo la contrastación e interpretar sus resultados.
- 7.- Estimar la pretensión de verdad de las conjeturas y las técnicas y formular los nuevos problemas originados por la investigación.

De ahí que un ciclo de investigación pueda medirse por los cambios que acarrea en nuestro cuerpo de conocimientos, o por los nuevos problemas que suscita. En la ciencia, la duda es creadora, estimula la investigación, la búsqueda de ideas que den razón a los hechos de un modo cada vez más adecuado. Desde Bacon, Descartes, Stuart Mill, etc., la metodología científica es capaz de suministrar indicaciones y proporcionar medios para evitar errores, pero no debe suplantar a la creación original, ni aun eludimos todas las equivocaciones.

Un método científico y la finalidad a la que se aplica constituyen la diferencia fundamental que existe entre la ciencia y la no ciencia; tanto el método como el objetivo científico son elementales para el logro intelectual. Quizá la única regla de oro del trabajo científico sería: "audacia en el conjeturar y rigurosa prudencia en el someter a contrastación las conjeturas". Por ello, el método científico es un rasgo característico de la ciencia, tanto de la pura como de la aplicada. Y puede afirmarse sin lugar a dudas que donde no hay método científico no hay ciencia. Pero el método no es infalible, ni autosuficiente; el método científico es falible y puede perfeccionarse mediante el análisis de los resultados, tanto directos como indirectos; el método tampoco es autosuficiente, no puede trabajarse sin hipótesis, con ausencia de conocimientos previos, pues se requieren estudios anteriores que puedan reajustarse y elaborarse. Todo método científico debe complementarse con métodos especiales que se adaptarán a las peculiaridades del objetivo científico.

En resumen, el método científico es la estrategia de la investigación científica; afecta al ciclo completo de la investigación, siendo independiente del tema en estudio; cualquier ciencia sería una disciplina que utilizaría el método científico con la finalidad de hallar leyes, (estructuras generales).

El blanco primario de la investigación científica es el progreso del conocimiento. El objetivo central de la investigación en las ciencias factuales sería mejorar el conocimiento del mundo de los hechos, y el de las ciencias aplicadas mejorar el control del hombre sobre los hechos. La diferencia entre ciencias formales y ciencias factuales sería que las primeras estudiaran ideas y las segundas, hechos. La ciencia formal sería autosuficiente, mientras que la ciencia factual dependería del hecho, por lo que hace al contenido, y del hecho experimental, para la convalidación; de ahí que los métodos sean medios arbitrados par alcanzar algunos fines.

El método científico serviría en primer lugar para incrementar nuestro conocimiento, para aumentar nuestro bienestar. Las ciencias puras persiguen un fin puramente cognitivo, mientras que las ciencias aplicadas (tecnológicas) utilizan el método general propio de la ciencia punta y varios métodos especiales que los aplica a fines prácticos. Las ciencias puras, que para algunos autores no se diferencian prácticamente en nada de las aplicadas, pretenderían mejorar nuestro dominio de las mismas. La investigación epidemiológica sería un claro ejemplo.

Para terminar, podríamos concluir que el objetivo central de la investigación en la ciencia factual pura sería mejorar nuestro conocimiento del mundo de los hechos; el de la investigación científica aplicada sería, mejorar el control del hombre sobre los hechos. "Lo que permite a la ciencia alcanzar su objetivo es su método". El método bien aplicado es capaz de descubrir las deficiencias científicas y corregir dichas deficiencias, es decir, construir representaciones de las estructuras intrínsecas que cada vez sean más adecuadas.

En el excelente texto que tenemos el honor de prologar se aplica la metodología científica al tipo de investigación epidemiológica. Rara vez es posible comprobar cómo de la teoría puede descenderse a la práctica con tanta sencillez y brillantez como en el presente texto. La claridad expositiva sería su característica más destacable.

En la primera parte se presentan los diversos tipos de investigación epidemiológica y sus aplicaciones clínicas. Su brevedad es muestra del profundo conocimiento de los autores y de su enorme poder de síntesis. El desarrollo es de fácil lectura. Se ha conseguido lo casi imposible: su comprensión.

La parte segunda de Análisis de Datos es quizá fundamental para el entendimiento de lo que es una metodología investigadora en el campo epidemiológico. La inferencia estadística, el error aleatorio, los tests en la comparación de grupos, los análisis multivariantes, etc. establecen la diferencia entre entendimiento completo del tema o su confusión. Lo mismo podríamos indicar de la parte tercera, que nos conduce a comprender que el soporte informático de la investigación está introducido en el mundo actual de manera prevalente. El análisis de las distintas fuentes de conocimiento señala la necesidad de su aplicación.

No quisiera dejar de comentar con sumo interés la parte cuarta, la que para el lector urólogo pueda ser quizá la más interesante, pues enseña como debe leerse críticamente cualquier tipo de publicación científica. Tras la valoración de los diversos tipos de comunicaciones, nos introduce en el campo de la crítica científica, hecho nada fácil cuando lo que tenemos delante es, supuestamente, el resultado de una investigación. De la lectura de estas deliciosas páginas debe inferirse nuestro análisis crítico en la lectura de cualquier tipo de comunicación; a ello se añaden comentarios acertados sobre las diferencias gramaticales y de estilo científico y literario o las cualidades del mismo; al final los autores dan 51 consejos útiles para evitar vicios y mal estilo en la redacción de cualquier comunicación científica.

Si como diría algún autor "hay libros buenos para siempre, libros buenos para un momento, libros malos para un momento y libros malos para siempre", sin la menor duda, el presente texto sería un libro bueno para siempre.

A todos los editores, autores y colaboradores nuestra felicitación y agradecimiento. Felicitación por el fruto de la obra bien hecha y gratitud, como científicos, por lanzar y defender la idea de que la metodología en las ciencias es un arma de la que nadie puede prescindir.

Creo que raramente se produce la feliz coincidencia de ideas claras, lenguaje preciso, solución a problemas y el planteamiento de dudas que nos abren horizontes para nuevas investigaciones.

Luis Resel Estevez
Catedrático de Urología
Universidad Complutense de Madrid
Presidente de la Asociación Española de Urología

PRÓLOGO A LA EDICION 1996

Cuando recibí la amable invitación del Dr. Rafael Burgos, para prologar la segunda edición de *Metodología de Investigación y Escritura Científica en Clínica* tuve conciencia que tal encargo era un medio para expresar las prolongadas y fecundas relaciones de los autores, a través de su labor en la Escuela Andaluza de Salud Pública, con un colectivo de profesionales sanitarios de dentro y fuera de Andalucía.

El presente libro en su primera edición (en la que hasta el prólogo era estimulante) ha constituido para muchos clínicos una ayuda de extraordinaria utilidad, bien como libro de texto en algunas de las actividades académicas impartidas por la EASP, bien como obra de consulta y orientación a la hora de elaborar proyectos, llevar a cabo estudios experimentales, analizar racionalmente los datos, interpretar literatura y expresar mediante lenguaje científico los resultados de nuestros estudios. Por todo ello, para los que apreciamos la obra, es un excelente noticia la aparición de una segunda edición, corregida y aumentada en su III^a sección (referente al soporte informático del investigador) y con una particular proyección a la Oncología, razón adicional para que se me encargara el presente comentario, en el que quiero expresar mis opiniones en torno al contenido de la obra, a su mensaje y al entorno de la cual procede.

La investigación científica distingue dos enfoques principales: inductivo y deductivo. El análisis mediante el que se disocian las partes de una idea, hasta el máximo grado posible para hacer accesibles a la mente humana conocimientos parciales y -de su elaboración científica y causal- inferir verdades generales.

El método científico que permite el progreso en las ciencias biológicas es el método analítico, que Descartes lo redujo a cuatro reglas aún actuales:

- * Admitir las verdades con extraordinaria precaución, cuando no haya nada que se oponga a ellas, aunque aparezcan con aparente y meridiana claridad
- * Desagregar los problemas en tantas partes como sea posible
- * Ordenar la pauta de adquisición de datos comenzando por los mas sencillos y terminando por los mas complejos
- * No omitir ninguna parte del problema en análisis de los datos, ni por otra parte, añadir parte extraña por conveniencia del éxito aparente.

Estos principios quedan reducidos en la investigación de cualquier hecho biológico (y de otra naturaleza) en la elaboración intelectual y ordenada de un programa metódico que consta de: observación de los hechos, hipótesis de trabajo y comprobación experimental. Por muy distintos que conceptual o físicamente parezcan los hechos investigados, el método se reduce a cumplimentar de manera rigurosa y adecuada a lo investigado, el programa descrito.

Durante la mayor parte de tiempo en que el conocimiento científico forma parte de nuestra cultura, el diseño experimental ha influido de manera decisiva (y adversa) en algunas ramas del saber biológico. Razones éticas, legales y sociales se han opuesto a que el ser humano vivo sea sujeto experimental y por ello la divergencia en el desarrollo de la denominada investigación básica médica, y la investigación clínica o epidemiológica, pero esas razones son ajenas al método científico en sí.

No ha sido sino hasta el siglo XX cuando se ha roto la restricción mental que constreñía la medicina clínica a una práctica denominada por Diego Gracia, medicina basada en la intención, en la que por la naturaleza intrínseca del ser investigado, el hombre, era ética y legalmente impensable ejecutar cualquier acto médico cuya finalidad no fuera la curación de la enfermedad, o su paliación. Moralmente era reprobable que el ser humano participara en una investigación cuya finalidad fuese la adquisición de conocimiento biológico, sin que tal saber fuera, forzosamente dirigido a curar, o a aliviar.

Frente a esa actitud, a lo largo de este siglo se ha ido imponiendo la necesidad de que la práctica médica tenga base científica, es decir, frente a la medicina basada en la intención, la medicina basada en la evidencia. Esta práctica médica, a la cual todos aspiramos, se sustenta en un método de investigación, que en nada difiere del que se ha aplicado en la denominada investigación básica. Lo que hoy es una manera "natural" de pensar en la ciencia médica se ha producido gracias al extraordinario desarrollo del método científico experimental en la clínica durante los últimos ochenta años, de modo que en contraste con la restricción "moral" que solo permitía una medicina intencional, hoy somos conscientes que toda labor clínica no basada en la evidencia, es decir en los resultados de la investigación científica, es éticamente reprobable.

A pesar del dramático cambio en nuestra concepción clínica, solo un 20% de nuestra actividad médica se basa en evidencias, pero ello sólo es un aliciente adicional para estimular los progresos en la investigación clínica; el enorme espacio ocupado por la práctica médica intencional es el campo de actuación presente y futuro de todo médico que tenga conciencia crítica de sus limitados conocimientos.

La profunda transición intelectual del médico a lo largo del siglo XX, muestra que si bien es una transformación general de la medicina, hay algunas especialidades médicas, como la oncología, en las que el pensamiento científico se ha impuesto. En efecto, el desarrollo de la especialidad parte de los años setenta, fecha en la que la medicina basada en la evidencia ya era "dogma" científico. Así mismo desde los inicios de la oncología, los pioneros de la especialidad se esforzaron por cuantificar y digitalizar cuando era posible, todas las variables asociadas a la enfermedad (evolución, extensión, terapéutica) y al enfermo (pronóstico) objeto de la especialidad, de forma que los datos que recolectamos pudieron ser sometidos fácilmente a procedimientos matemáticos. El método de investigación (relación causa-efecto) tiene su soporte más importante en la estadística, ciencia que es una especialidad matemática.

El presente libro permite a cualquier profesional con intención de iniciar una investigación epidemiológica entender las bases operativas del proceso intelectual de su trabajo. En todo caso y aun en ausencia de proyecto formal de investigación, proporciona la formación adecuada que un médico requiere en estos momentos para desarrollar su sentido crítico en la práctica clínica. Tales virtualidades derivan de su contenido y de su estilo.

La primera parte describe y conceptualiza los diversos diseños, entendidos como herramientas de trabajo, que tienen aplicación en los problemas que acomete la investigación clínica.

La segunda parte se dedica a los métodos estadísticos de análisis de datos, con una particular y gratificante claridad. Los médicos, que por desgracia carecemos de una sólida formación matemática, tenemos cierta dificultad en la comprensión de ciertos conceptos matemáticos; en consecuencia se aprecia doblemente el esfuerzo de simplificación de sus autores.

La tercera parte se dedica al soporte informático aplicado a la investigación clínica. Se ha corregido y modernizado, haciendo algo más fácil la utilización de nuestros PCs

La cuarta parte se dedica a las normas de la escritura científica y la lectura crítica de artículos originales; constituye una fuente de información espléndida sobre aquellas cuestiones que nadie nos enseñó a utilizar y que sólo a través de errores personales vamos perfeccionando. Se podría decir que es tan útil como oportuno y que viene a representar algo similar a lo que W. Allen titulaba: "Todo aquello que Vd. quería saber sobre....."

Este texto no es "uno más" de entre los que se utilizan en la educación básica del investigador. Los autores, voluntaria o involuntariamente, nos permiten extraer una conclusión fundamental de la atenta lectura de su libro: la INVESTIGACIÓN, con mayúsculas, no es patrimonio de un grupo selecto de profesionales, que imbuidos de un aura especial parecen pertenecer a la reducida plantilla de los intelectuales con derecho a pensar; cualquier médico puede aplicar el método científico a su labor. Naturalmente, nadie debe pensar que el progreso en la investigación clínica solo derivará de aportaciones individuales geniales. El progreso en esta actividad humana, como en muchas otras viene de la mano de la cooperación, más que de la competición, pero en esta colaboración nadie está excluido, a condición de poseer una mente sana y educada en ejercer el método científico, y a esto último contribuye esta obra. Mostrarnos el mundo de la investigación como una actividad factible es una importante aportación de este libro. Todos podemos colaborar. Karl Popper escribió: "El espíritu de la Ciencia ha cambiado a consecuencia de la investigación organizada. Esperemos que a pesar de esto, siempre haya individualidades. Ninguno de nosotros debería renunciar a ejercer esa aportación individual".

No se puede finalizar este comentario sin hacer referencia a sus autores y a su ámbito de trabajo. Prácticamente todos ellos han sido mis profesores a lo largo de diferentes cursos que he recibido en la EASP; con algunos guardo antiguas y estrechas relaciones. No podría olvidar que Emilio Sánchez-Cantalejo fue el responsable del tratamiento estadístico de los datos que conformaban la investigación de mi Tesis Doctoral allá en el año 1974. Todos son sin duda representantes del "espíritu de la EASP". La Escuela Andaluza de Salud Pública ha sido un faro intelectual de difusión de un estilo de hacer una práctica médica firmemente comprometida con la sociedad en la que nace. Yo, que pertenezco a la institu-

ción universitaria, tengo que lamentar que en algunos Departamentos (al menos de los que yo conozco en el ámbito de las ciencias médicas) aun impere una noción elitista de "excelencia" intrínseca que los convierte en "miniuniversos-islas", en los que la separación de la realidad social (de cuyo entorno deriva su financiación) los acerca a la concepción de Universidad y del saber propia del mundo de Guillermo de Baskerville (Umberto Eco), situación que se agrava por las dificultades en el ejercicio de la crítica y la autocrítica. En este panorama, ha sido muy importante que una institución pública dedicada a la formación continuada en investigación epidemiológica, práctica clínica y gestión entienda que no se puede estar de espaldas a la realidad, y que nuestro deber es construir la alternativa sanitaria que la sociedad demanda. Bajo esta óptica resulta particularmente coherente y apreciable la labor de este grupo de compañeros que se ha plasmado en la publicación que he tenido oportunidad de comentar.

Granada, Junio de 1996

Dr. José L. García Puche
Jefe del Servicio de Radioterapia y Oncología.
Hospital Clínico San Cecilio
Profesor Titular de la Universidad de Granada

AGRADECIMIENTOS PRIMERA EDICIÓN

Son muchas las personas que de una manera u otra han contribuido a la realización de esta monografía. Los autores quieren agradecer especialmente a Víctor Baena, José Carnero, Aurelio Cayuela, Alberto Fernández, M^o Ángeles Fernández, Carmen Martínez, Rosa Miñarro, Víctor Sarmiento y a José María del Rosal la lectura de algunas partes del manuscrito y sus interesantes comentarios.

A José Manuel García y Juan Ortega por su apoyo en una parte importante del diseño de los gráficos. A Giulia Fernández y Ana M^a Galán por su eficiente labor de secretaría.

Finalmente, quisiéramos agradecer a SCHERING PLOUGH por el patrocinio de esta obra.

Los Autores

INTRODUCCIÓN A LA EDICIÓN 1996

La primera edición de Metodología de Investigación y Escritura Científica en Clínica ha tenido una entusiasta acogida por parte de médicos clínicos, tanto de médicos especialistas como de médicos de atención primaria. También ha rebasado con creces el objetivo original de servir de libro de texto para alumnos de cursos de post-grado en las disciplinas de epidemiología, bioestadística, metodología de investigación y redacción científica en la Escuela Andaluza de Salud Pública y otras instituciones de enseñanza superior especializadas en las disciplinas mencionadas.

Agotada la primera edición al cumplirse un año y medio desde su aparición, la EASP se complace en presentar esta edición 1996 la cual reproduce íntegramente la 1ª edición con algunos cambios y adiciones menores necesarios que se describen a continuación.

La Parte I, "Diseños de investigación epidemiológica y sus aplicaciones en Clínica" reproduce la primera edición, con algunas correcciones en el texto tendientes a aclarar y realzar su contenido.

La Parte II, "Análisis de datos en investigación clínica y epidemiología" conserva también el mismo texto, tablas y gráficos de la edición anterior, incluyendo algunos ajustes de precisión conceptual y terminológicos.

La Parte III, "El soporte informático de la investigación" ha sido objeto de una minuciosa revisión; esto ha significado un notable enriquecimiento del texto y la adición útil de media docena de tablas y figuras las cuales ilustran y completan el texto.

La Parte IV, "Escritura y lectura crítica de artículos científicos" conserva el texto original de la primera edición, a lo cual se ha agregado un capítulo breve pero fundamental que relaciona las etapas del proceso de investigación con la estructura del artículo científico original. Además se han revisado extensamente las secciones sobre tablas y gráficos, haciendo referencia cruzada a algunas tablas y gráficos presentados en las Partes I y II de la obra.

El libro conserva los Apéndices originales, a lo cual se han agregado un apéndice sobre revisión bibliográfica, otro sobre ejemplos de tablas y otro sobre ejemplos de gráficos, todos de gran valor instruccional y de aplicación práctica para el lector.

La Escuela Andaluza de Salud Pública se complace en presentar la edición 1996 de esta importante obra, única en su género; cumple de esta manera con su misión de difundir conocimientos, desarrollar habilidades y promover cambios de actitudes en los profesionales de los servicios de salud en su delicado y complejo quehacer diario.

La EASP agradece a Shering Plough la renovación de su patrocinio de esta importante obra.

*Manuel Alejandro Bobenrieth Astete
Profesor de escritura científica
Coordinador Unidad Técnica de Publicaciones
Escuela Andaluza de Salud Pública
Granada, Agosto de 1996*

INTRODUCCIÓN A LA PRIMERA EDICIÓN

La transmisión de información

El conocimiento humano se ha mantenido desde el albor de los tiempos gracias a la transmisión de información, la que comenzó por ser de forma oral hasta la invención de la escritura en el siglo XXVIII antes de J.C.. La primicia de esta invención la reclaman los jeroglíficos egipcios, la escritura china, la escritura cuneiforme de los asirios, los jeroglíficos aztecas y la escritura maya.

La primera forma conocida de comunicación escrita entre pensadores e investigadores es la correspondencia personal mediante la *Carta Erudita* del siglo XVII, a la cual se agregan los catálogos de libros anuales, los noticieros, los periódicos y las primeras revistas científicas generales que se inician en el siglo antes mencionado. Las primeras publicaciones a las que puede aplicarse este calificativo aparecieron en 1665: las *Philosophical Transactions* de la Regia Sociedad Londinense y el *Journal de Scavans* de Denis de Sallo. En 1670, inició su publicación la *Miscellanea curiosa* de la Academia Naturae Curiosorum que, como órgano de una "Gesellschaft Naturforschender Aertze", tenía un contenido fundamentalmente médico, complementado con temas de historia natural. Recuérdese asimismo la fecha fundacional del *Acta médica et philosophica hafniensia* en 1671 de Thomas Bartholin y la del *Journal des nouvelles découvertes sur toutes les parties de la médecine* en 1679 de Nicolas de Blegny, esta última la primera revista médica en idioma vulgar.

Al principio las revistas intentaron únicamente ofrecer de manera puntual resúmenes de los nuevos libros científicos que iban apareciendo; gradualmente esta función fue pasando a segundo plano para convertirse en medios de comunicación de las investigaciones y descubrimientos más recientes publicados en ningún libro.

El siglo XVIII ve nacer las primeras revistas científicas especializadas y algunas publicaciones de sociedades científicas. Sin embargo esta comunicación científica escrita derivada de la investigación que comenzó hace aproximadamente tres siglos, dio a la luz el artículo científico original tal y como lo conocemos hoy día hace apenas 100 años.

El crecimiento, tanto del número de revistas como del número de artículos originales publicados, ha sido exponencial desde 50 revistas en 1750 con aproximadamente igual número de artículos en ese año hasta 23.480 revistas biomédicas y de salud en 1994 con unos 500.000 artículos publicados por año. El ritmo de este crecimiento es mucho más rápido que el de la mayoría de los fenómenos sociales. Mientras que la población se duplica cada cincuenta años y la renta de los países desarrollados cada veinte, el tamaño de la información científica se convierte en el doble solamente en diez años si utilizamos una concepción amplia de ciencia, o en quince si la definimos de modo más exigente. Cada quince años, por ejemplo, se duplica el número de revistas científicas y de "abstracts" o resúmenes de todos los campos de la ciencia pura y aplicada, lo mismo que el de compuestos químicos conocidos o el de miembros de instituciones científicas.

Un primer corolario de este crecimiento exponencial es la "contemporaneidad" característica de la ciencia moderna. Llamando coeficiente de contemporaneidad al porcentaje que significa la ciencia actual respecto del conjunto de la de todas las épocas, resulta que al período de duplicación de diez años corresponde un coeficiente en torno al 96% y al de quince años otro alrededor del 87,5%. De esta forma, el número de artículos sobre rayos X, geometría euclidiana o psicología experimental publicados en dicho período, es igualmente el 96% del total publicado hasta ahora.

Otra consecuencia del carácter del conocimiento científico moderno y de la dinámica de la comunicación entre los grupos que lo cultivan es el rápido envejecimiento de la información científica, es decir, la acelerada tendencia a que caiga en desuso (obsolescencia). Este fenómeno no fue, sin embargo, de aparición inmediata como en el caso del crecimiento exponencial sino que, tras una serie de etapas intermedias, no alcanzó plena vigencia hasta mediados del siglo XIX.

El proceso de investigación en sus orígenes consistía en la mera observación de fenómenos de los cuales se extraían conclusiones; sin embargo a nivel médico hacia la mitad del siglo XIX en que Louis Pasteur confirmó la teoría microbiana de las enfermedades y elaboró métodos de cultivo, se hizo necesario garantizar la reproducibilidad de los experimentos para así acallar a los partidarios de la generación espontánea. Pero aunque la observación y la reproducibilidad son requisitos de la investigación, con el devenir de los años ha surgido la aplicación del "razonamiento inductivo" y por tanto la necesidad de efectuar múltiples observaciones particulares con rigor metodológico de las que se pueda extraer conclusiones generales. Esta cadena de sucesos que presentamos en forma tan resumida conlleva una cantidad de contaminaciones que el investigador debe controlar no solo a lo largo del proceso de investigación, sino también en la de comunicación de la misma; de nada servirá un buen trabajo de investigación si no se difunde y transmite el conocimiento adquirido en forma correcta y

oportuna. Igualmente se debe evitar el publicar artículos que no aporten nada, circunstancia bien planteada desde finales del siglo XIX por George Eliot, ensayista y novelista británica cuando escribía "Bendito es el hombre quien no teniendo nada que decir se abstiene de evidenciar este hecho con palabras".

El conocimiento científico

El quehacer clínico tiene que ver con la aplicación del conocimiento científico al diagnóstico y tratamiento de problemas concretos de salud y de enfermedades en individuos y en grupos. El conocimiento científico es la consecuencia del pensamiento científico el cual se reconoce por su objetividad, su racionalidad y su sistematicidad. Hoy en día existe consenso sobre una docena de características del conocimiento científico, cuya descripción sucinta sería como sigue:

Fáctico: Parte de hechos de la realidad que se generan, de alguna forma, en la experiencia sensible;

Trascendente: Parte de los hechos pero va más allá de los mismos ya que se eleva por encima de un nivel o límite dado;

Metódico: Procede en forma organizada y planea lo que persigue averiguar, la forma de obtenerlo, de analizarlo e interpretarlo. Para ello dispone de procedimientos tales como la inducción, la deducción y la analogía;

Analítico: Distingue y separa las partes de un todo hasta llegar a conocer sus elementos, las relaciones entre éstos y los principios que los rigen;

Claro y preciso: Los conceptos científicos y los problemas objeto de estudio se definen sin vaguedad, sin ambigüedad y sin confusión;

Simbólico: Utiliza el lenguaje cotidiano, pero va más allá de éste. Crea su propio lenguaje con palabras, signos y símbolos que suelen someterse a reglas;

Verificable: Los resultados y conclusiones deben someterse a prueba mediante observación y -de ser posible- a experimentación;

Explicativo: A diferencia del conocimiento cotidiano, el conocimiento científico no se conforma únicamente con los hechos tal como se dan; investiga su asociación y -de ser posible- sus causas. Procura explicar los hechos en términos de teorías, de leyes y de principios;

Predictivo: Intenta explicar lo presente y lo pasado; además aventura el futuro. La capacidad de predecir le sirve al científico para modificar acontecimientos mediante ciertas intervenciones en beneficio de la sociedad. (Pronósticos médicos, predicciones meteorológicas, etc.);

Abierto (Escéptico): Noción de que cualquier proposición o afirmación, aunque provenga de autoridades importantes, es susceptible de duda, de cuestionamiento y de controversia;

Comunicable: Sujeto de transmisión oral o escrita basada en datos sobre hechos observables, presentación, análisis e interpretación de los mismos.

Util: Es aprovechable, beneficioso, conveniente, fructífero, sensible y hasta lucrativo. Nuestro mundo actual sin la ciencia inmersa en él retornaría a la época de las cavernas.

El conocimiento científico no está exento de limitaciones, las cuales se pueden agrupar en cuatro categorías:

- a) **Problemas de complejidad:** Los asuntos que son objeto del conocimiento científico suelen ser de difícil comprensión por cuanto se componen de muchos elementos muy diversos e interrelacionados; Identificar sus características o atributos (variables) y establecer la asociación entre ellos (casual o no), resulta en general difícil y -a veces- imposible.
- b) **Problemas de medición:** La ciencia es inconcebible en ausencia de algún tipo de medición. Muchos avances científicos están condicionados en gran medida por los procedimientos e instrumentos de medición empleados. La medición se halla en una posición estratégica entre la teoría y la experiencia; entre la razón y la naturaleza.
- c) **Problemas de control:** Inherente a cualquier tipo de diseño de investigación está la necesidad de control, en términos de comparación o comprobación. Los tipos de control más frecuentemente utilizados son: control por manipulación, control estadístico y control por aleatorización.
- d) **Problemas de ética:** La ética, parte de la filosofía que estudia la moral y su repercusión en los actos humanos, tiene por objeto los juicios de valor cuando se aplican a la distinción entre el bien y el mal, lo permitido y lo prohibido en la investigación. Los problemas de ética tienden a limitar los estudios experimentales, sea por acción o por omisión.

La metodología de investigación

La metodología es el conjunto de métodos que se siguen en una investigación científica, la cual se funda en principios y se materializa en procesos sistemáticos los cuales se expresan en actividades. Cualquier sistema de conocimiento se basa en el empleo de métodos apropiados; el sistema de conocimientos en clínica no hace excepción.

Una revisión atenta de los diversos libros de texto y de consulta sobre metodología de investigación muestra -más allá de la diversidad de estructuras y de nomenclatura- que el método científico se desarrolla en 3 grandes etapas consecutivas: etapa conceptual o etapa teórica; etapa empírica o etapa práctica y etapa interpretativa o etapa del significado.

Cada una de las etapas consta de diversos componentes los cuales pueden esquematizarse como sigue:

1. Etapa conceptual (teórica). Observación de hechos; identificación del problema general; revisión bibliográfica (antecedentes del problema); marco teórico; definición del problema de investigación; definición operacional de términos; definición de variables; importancia y limitaciones del estudio; definición de objetivos; formulación de hipótesis.
2. Etapa empírica (práctica). Diseño, el cual consiste en el enfoque, la estrategia de abordaje del problema de investigación para el logro de los objetivos definidos al final de la etapa conceptual. El autor debe justificar la selección del diseño en particular e incluir el plan de la investigación. El diseño incluye los procedimientos, los métodos y las técnicas para la selección de los sujetos, la recolección de la información y el análisis de los resultados. El diseño representa la conexión entre la hipótesis y los datos, entre las ideas y la realidad. El diseño varía según la naturaleza del problema y el tipo de investigación.

La decisión sobre el tipo de diseño acarrea una serie de otras decisiones en la ejecución de la investigación. De entre estas se destacan: universo; muestra; unidad de observación; unidad de medida; información necesaria; fuentes de información; recolección de los datos; tabulación de los datos; presentación de la información; análisis de la información.

3. Etapa interpretativa (significado). Convalidación de los datos, es decir la solidez de los mismos en razón de su origen, del método de recolección y de la firmeza estadística; validez interna; validez externa; relación de los hallazgos con los objetivos e hipótesis; relación de los hallazgos con los hallazgos publicados por otros autores; aspectos nuevos e importantes del estudio.

El escenario de las publicaciones

En los últimos 30 años la cantidad de artículos científicos publicados ha experimentado una importante eclosión, siendo materialmente imposible que un médico de cualquier especialidad pueda efectuar una lectura detallada de cada uno de los artículos publicados debido a serias dificultades relativas a su disponibilidad, al tiempo requerido y a su coste. Por otra parte en las publicaciones médicas en general, y las urológicas en particular, hay cada día una mayor proliferación de términos epidemiológicos y estadísticos. Además existe gran desconocimiento sobre la metodología de escritura científica tanto por los lectores como por los autores; incluso en ocasiones hasta por los editores de revistas científicas.

Estas circunstancias se traducen en serias dificultades para la comprensión de la lectura de los artículos por: a) incompreensión o tergiversación de los términos estadísticos y epidemiológicos; b) faltas de estructura y de estilo de los artículos; c) apabullamiento ante el tipo de información presentada. En el caso de asistentes a congresos o jornadas se inhibe en mayor o menor grado el debate lo cual afecta negativamente la crítica constructiva que es tan buena aliada del conocimiento científico.

Estos hechos generan tres posturas claramente definidas en aquellos clínicos no versados en términos estadístico-epidemiológicos de reciente implantación:

- a) El rechazo de la epidemiología y de la estadística, argumentando "Bueno, dejémonos de estadística y vayamos a lo importante"; ésto es lamentable ya que se privan de herramientas indispensables para el análisis y la inferencia de conclusiones a grupos de mayor ámbito, lo cual permitiría el conocimiento del estado de salud de la población con respecto a determinadas enfermedades.
- b) La elevación de los términos estadísticos a la categoría de verdad absoluta hablando con ocasión y sin ella de "p", de significación estadística, de chi-cuadrado, de t-Student, etc. De esta manera se basa la validez de un trabajo de investigación exclusivamente en la significación estadística la cual -como es bien sabido- no es necesariamente equivalente a la significación clínica. Como era de esperar y como aconseja el sentido común, ninguna de las dos posturas es acertada.
- c) La aceptación del valor de los conceptos estadísticos y epidemiológicos en la investigación clínica y la decisión de conocerlos y comprenderlos un poco más para utilizarlos mejor. Este libro apunta a fortalecer esta postura sensata e inteligente.

Sin duda, para una correcta interpretación de los artículos con determinaciones estadísticas y diseños epidemiológicos o para la elaboración de los mismos, es necesario un adecuado conocimiento de la estadística y de la epidemiología, además de un adecuado conocimiento de la metodología de escritura científica. Así es importante conocer el alcance de los términos significación estadística, error α , error β , sensibilidad, especificidad, valor predictivo, hipótesis nula, métodos paramétricos o no paramétricos, etc....., así como su correcta interpretación. Por otra parte es importante distinguir entre medidas de asociación y de magnitud, conocer cual es el impacto poblacional, la importancia del muestreo probabilístico, las diferencias entre los diversos diseños de estudios epidemiológicos y cual es el más adecuado al objetivo clínico definido.

De la misma manera deberíamos saber los requisitos de un buen título y del buen resumen de un artículo original; de qué elementos debe constar la introducción; qué le debemos pedir al material y métodos a la luz del diseño y de los objetivos del

artículo; cómo debemos presentar los resultados; cómo conducir la discusión y elaborar las conclusiones; cómo citar correctamente las referencias bibliográficas.

En resumen, la estadística y la epidemiología constituyen herramientas fundamentalmente de investigación, las cuales dan validez y peso específico a las investigaciones. Estas armas nos ponen en guardia frente a estudios que tienen una formulación inadecuada e insuficiente de los objetivos; estudios cuyos resultados aparecen como estadísticamente significativos pero que tienen errores de diseño que los invalidan. Una vez asegurada la corrección en los aspectos claves de contenido se debe garantizar su presentación y publicación correcta, clara y concisa a la comunidad científica.

La génesis del libro

Tras lo expuesto en los apartados anteriores parece clara la necesidad de un libro como el que presentamos, no tanto para obtener una completa formación la cual sin duda requiere algo más que un libro de estas características, sino para aclarar las lagunas sobre metodología y servir de ayuda en la resolución de problemas prácticos; además para la introducción y comprensión de un lenguaje que a muchos clínicos les resulta abigarrado y confuso.

Parece necesario explicar porqué un libro sobre metodología de investigación y publicación científica en clínica contiene tantos ejemplos urológicos. La obra tiene su origen en el ámbito de la urología desde el momento en que los urólogos editores de este libro recibieran formación en epidemiología, estadística y escritura científica en la Escuela Andaluza de Salud Pública.

En efecto se ha gestado un microentorno metodológico alrededor del Servicio de Urología del Hospital Carlos Haya de Málaga, dónde como es de conocimiento de los urólogos del país, es usual plantearse muchas de las cuestiones que ahora se plasman en este libro. Estamos convencidos que los urólogos no pueden contentarse sólo con sus conocimientos especializados sino que deben inquietarse por temas metodológicos y emplear algún tiempo libre en procurarse formación en éstas áreas.

Estas inquietudes -que fueron al principio la iniciativa de dos profesionales- quedaron afianzadas cuando en el LVIII Congreso de la Asociación Española de Urología celebrado en Málaga en 1993, la asamblea confirmó la propuesta de una mesa redonda sobre "*Importancia de la Metodología de investigación en la comunicación científica urológica*" que un año más tarde se desarrollaría en Sitges en el seno del LIX Congreso y en la que participaron además otros dos urólogos con formación específica, concretamente los Dres Carnero Bueno y del Rosal Samaniego.

Simultáneamente al desarrollo de la mesa redonda y a raíz de una conversación de los dos urólogos editores nació la idea sobre la realización de este libro. Esta idea se fue depurando con el tiempo hasta contar con la acogida entusiasta de la Escuela Andaluza de Salud Pública la cual se materializó en la cooperación como autores de cuatro profesores de las asignaturas de epidemiología, bioestadística, informática y redacción científica. Además se incluyó la obra dentro de la Serie de Monografías de la EASP, lo cual asegura la recuperación posterior de la información que contiene mediante su inclusión en el Index Medicus y otras bases de datos nacionales y europeas.

La estructura del libro

El libro se presenta estructurado en cuatro partes las cuales contienen 27 capítulos con 107 secciones, siguiendo -en base selectiva-la secuencia del proceso de investigación y de la publicación de su informe en la forma de un artículo científico original.

La PARTE I toma el proceso de investigación al momento del diseño, es decir, de la decisión sobre el enfoque, el abordaje y el plan de la investigación. Esta decisión sucede al término de la etapa conceptual y comanda -de hecho- la fase empírica de la investigación. El autor describe a lo largo de 7 capítulos los tipos de estudio, los errores más frecuentes en los diseños, algunas formas de control y las principales medidas de frecuencia. Luego progresa hacia la explicación clara de los tipos de estudio más frecuentes: estudios descriptivos, estudios de seguimiento, estudios de casos y controles, estudios cuasi experimentales y estudios aleatorios controlados.

La PARTE II retoma el proceso donde lo dejó el diseño ya ejecutado y se centra en aspectos claves del análisis de datos en investigación clínica y epidemiológica. Aquí los autores se explayan en seis capítulos sobre temas tan importantes como el análisis descriptivo, los tests diagnósticos, la inferencia estadística, los tests para comparar dos grupos, el análisis multivariante y el análisis de decisión clínica.

La PARTE III se centra en el soporte informático en la investigación con una introducción y dos capítulos los cuales tratan sobre información y datos, y el software disponible.

La PARTE IV cierra la obra con cuatro capítulos dedicados a la escritura de artículos científicos y la lectura crítica de aquellos artículos ya publicados. Después de una breve introducción sobre formas de comunicación científica el autor se concentra en la "anatomía", la "fisiología" y la "fisiopatología" del artículo científico original. Se incluye además una reseña breve sobre las partes del artículo de revisión, para cerrar con importantes anotaciones prácticas sobre el estilo en la comunicación cien-

tífica. Esta parte IV agrega tres importantes apéndices, con ejercicios sobre cómo redactar títulos correctos; siete tipos de pautas para evaluar artículos originales; los requisitos uniformes para preparar manuscritos enviados a revistas biomédicas.

Cada una de las partes se apoya en referencias citadas en el texto, para efecto de verificación y profundización.

El conjunto de las cuatro partes configura una obra con un contenido global metodológico, estadístico y estilístico único en su género, de trascendencia y de utilidad.

Los objetivos y la audiencia de la obra

Es cada día más frecuente el constatar que muchos médicos clínicos tienen excelentes ideas de investigación pero que generalmente carecen de las herramientas conceptuales y metodológicas para implementarlas. Estas herramientas consisten en habilidades para el diseño de la investigación, planificación de la investigación, implementación del plan, presentación, análisis e interpretación de los datos y para la redacción clara y concisa del informe. Hoy existen varios libros de investigación que incluyen capítulos sobre diseño y sobre tratamiento estadístico de los datos; sin embargo el problema consiste en que los profesionales de los campos de la clínica -que actúan como investigadores ocasionales- rara vez tienen la formación necesaria para ubicar tantas fuentes y valorar aquella información que les ayude resolver un problema específico.

El propósito básico de este libro es actualizar al médico clínico sobre algunos aspectos claves del proceso de investigación y de la escritura de su informe. No ha sido la intención de los autores cubrir todas las áreas de la investigación en forma de un tratado o de un manual, sino simplemente ofrecer una panorámica rápida del cómo diseñar una investigación, cómo presentar y analizar los datos, cómo utilizar los recursos que nos provee la informática y cómo publicarla en la forma de un artículo científico.

El libro ha sido escrito con el objetivo de ayudar a resolver problemas prácticos que se le presentan al clínico al programar y realizar una investigación correcta. No intenta -en consecuencia- ser un libro sobre la teoría de la investigación ni de la filosofía de la ciencia, sino de enfoques y de técnicas concretas. Naturalmente que esta orientación práctica no excluye la necesidad de explicar -cada vez que el caso lo requiere- el "qué" y el "porqué" que justifican el "cómo" hacerlo mejor.

Otro objetivo de esta obra tan especial es ayudar a los clínicos a evaluar en forma más inteligente y más crítica los trabajos que se presentan a reuniones científicas y aquellos otros que se publican como artículos, aún cuando éstos no hagan investigaciones propias.

Estas reflexiones justifican que el destinatario principal del libro sea el médico clínico, tanto el médico general de atención primaria, como el médico especialista. Para ello se ha intentado escribirlo de forma comprensible y práctica. El libro va dirigido principalmente a médicos sin conocimientos previos de epidemiología, de estadística, de informática, ni de redacción científica.

Creemos que la obra puede ser igualmente útil para enfermeras, odontólogos, farmacéuticas, fisioterapeutas, veterinarios, salubristas y otros profesionales y técnicos del campo de la salud. Igualmente estamos convencidos de que la monografía será de gran utilidad para estudiantes de medicina y de enfermería quienes tendrán la oportunidad de incorporar precozmente las metodologías de investigación y de publicación científica como valor de vida en su quehacer diario.

Cada autor asume plena responsabilidad de su capítulo y de sus secciones; además los autores han hecho el máximo esfuerzo para unificar criterios y coordinar los contenidos. Naturalmente ésto no se contrapone a la coexistencia pacífica y natural de estilos un tanto distintos, propios de obras de autores múltiples.

En suma, este libro intenta descubrir los secretos de ese no tan fácil arte de convertir el pensamiento científico sobre un problema real por resolver en un artículo científico publicado. Los autores esperan que la obra cumpla con el gran objetivo que se trazaron al rellenar cada página en blanco: ser útil.

Rafael Burgos Rodríguez
Málaga, octubre de 1994

PARTE I
DISEÑOS DE INVESTIGACION EPIDEMIOLÓGICA
Y SUS APLICACIONES EN CLÍNICA

Emilio Perea-Milla López
EDITOR

A. TIPOS DE ESTUDIO

E. Perea-Milla López

1. Conceptos generales

La investigación en clínica, cuando utiliza el enfoque epidemiológico, obedece a objetivos muy variados. El interés de un determinado estudio puede ir desde la simple descripción de una realidad, en una población o un conjunto de pacientes, hasta la verificación de una hipótesis en relación a factores de riesgo o factores pronósticos. No se trata de hacer una mera 'recogida de datos' que algún tipo de análisis estadístico pueda 'exprimir'. Por el contrario, existe una fase previa de elaboración de una estrategia para comprender el problema que se afronta, definir los grupos que se van a comparar y decidir cuál será el tratamiento o análisis al que se someterán los datos. La elección del diseño más adecuado ha de preceder siempre al proceso de recolección de información. Se trata de evitar que un trabajo, muchas veces arduo, quede invalidado por la falta de un diseño correcto, que ningún análisis estadístico puede salvar.

Un esquema explicativo de lo que la mayoría de los estudios o diseños pretenden, inspirado en Alhbohm y Norell ¹, es como sigue (Fig. I.1):

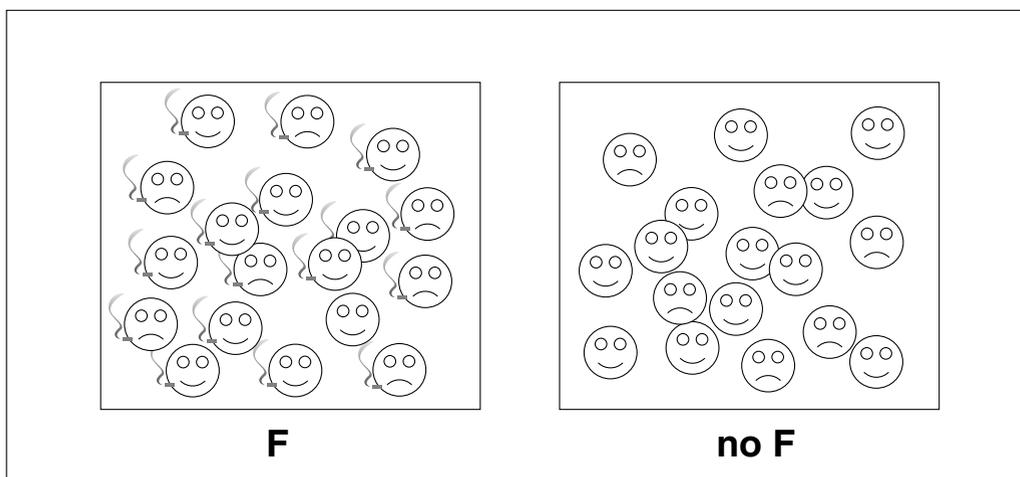


Fig. I.1- Estudios o diseños epidemiológicos: grupos y comparación de grupos.

Grupos. La primera idea implícita en la figura I.1 es que trataremos con **grupos** de personas. En el trabajo clínico diario se reúne información a través de la anamnesis, la exploración y las pruebas complementarias. Ese proceso, completado con un adecuado diagnóstico diferencial, lleva a la verificación o rechazo de distintas hipótesis sobre la situación de salud de un paciente, lo que denominamos *diagnóstico*. El enfoque epidemiológico, sin embargo, utiliza información sobre la distribución de los problemas de salud y sus determinantes en colectivos de personas, ya sean enfermas o sanas.

Comparación de grupos. Una segunda idea es que, si bien en muchos casos es suficiente la descripción de la situación en un colectivo, de alguna forma siempre se realizan **comparaciones** con lo que ocurre en otros. Esas comparaciones son a veces implícitas, intuitivas, pero en la mayor parte de los diseños habrá que explicitar un grupo de comparación, que además sea el más adecuado. Cuando el interés del estudio sea determinar si la presencia de un factor de exposición '*F*' o alguna característica (por ejemplo, nivel sérico elevado de antígeno prostático-específico, PSA) se asocia con la aparición de una enfermedad '*E*' (carcinoma de próstata, en este caso), será imprescindible conocer cómo se distribuye esa enfermedad en la gente que no presenta la característica (*noF*). Es más, no se podría afirmar nada sobre la posible relación si sólo se conociera qué ocurre en el grupo expuesto a '*F*'.

La distribución de la enfermedad. Los dos puntos anteriores (la información sobre grupos y la comparación entre ellos) se basan en una asunción previa: los problemas que afectan a la Salud tienen una distribución **no aleatoria**, es decir, no se distribuyen al azar. Como consecuencia, es posible encontrar diferencias entre los distintos grupos que se comparen, que dependerán de la desigual distribución entre ellos de distintos factores de exposición; lo mismo podría afirmarse sobre los factores pronósticos. Si no fuera así no tendría sentido comparación alguna.

2. Secuencia del proceso

El abanico de posibles *diseños de estudio* es muy amplio, y la elección de un determinado *tipo de estudio* estará condicionada, entre otras, por la pregunta de investigación planteada, en la forma de objetivos o hipótesis. Básicamente, el trabajo con datos procedentes de grupos de personas (sean enfermas o sanas), trata de responder a alguna de las siguientes dos preguntas²:

- 1) ¿Cuál es el comportamiento de un problema que afecta a la Salud?

Aún cuando suele identificarse *problema de salud* con *enfermedad*, el objetivo de la investigación puede referirse a muerte, recurrencia, actitudes, curación o cual-

quier otro efecto relacionado con la situación de salud de las personas. Notaremos en adelante la presencia de cualquier efecto como *E*, y como *no E* su ausencia. La respuesta a esta primera pregunta se encuentra mediante la descripción de variables relacionadas con las características de las personas, el lugar o el tiempo en los que se presenta el *problema*.

2) ¿Cómo se explica que ese comportamiento sea así y no de otra manera?

La respuesta intenta analizar qué factores están relacionados con el problema en estudio, en qué medida y qué cantidad de daño a la salud es atribuible a la presencia o ausencia de esos factores. Notaremos a los factores predictores de cualquier efecto mediante *F*, y su ausencia como *no F*.

Además, debe tenerse en mente el estado del conocimiento previo sobre el problema a investigar (Fig. I.2). Este será indudablemente un determinante del tipo de pregunta que se plantee y, por tanto, del tipo de estudio.

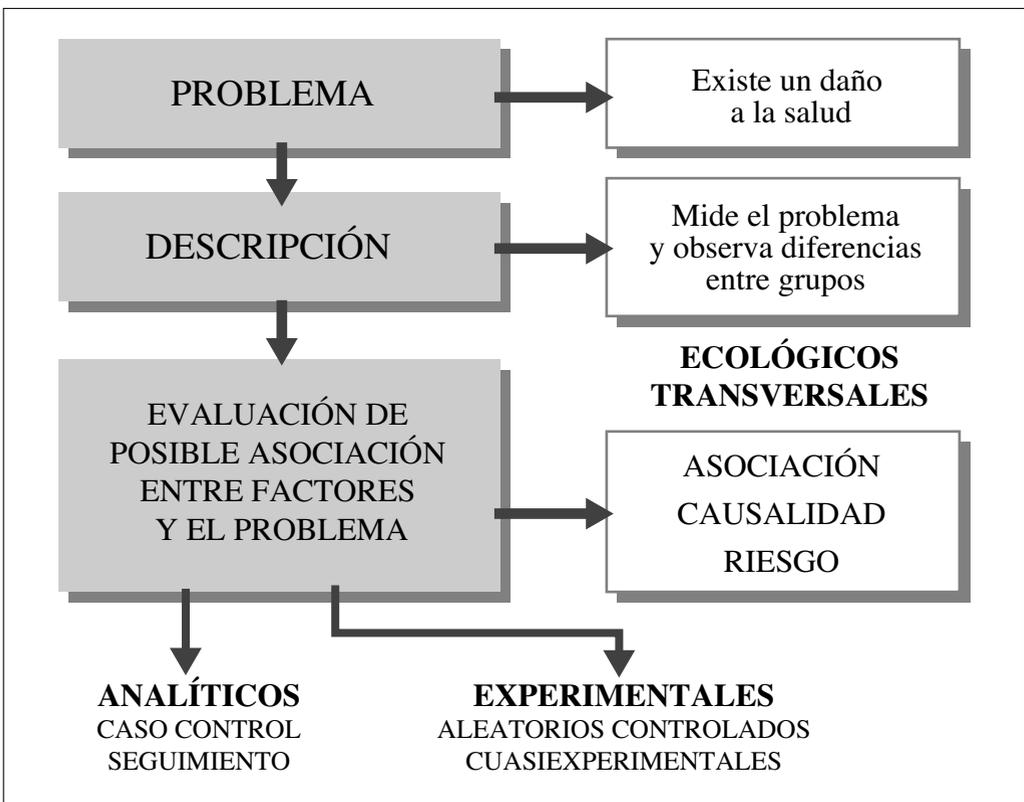


Fig. I.2- El tipo de diseño. Sus indicaciones según el conocimiento previo sobre el problema.

Así, es posible que nos encontremos en una primera fase, en la que lo importante es definir precisamente en qué consiste el problema. Una situación semejante a la que se planteó cuando los médicos de un hospital de Madrid en la década de los ochenta detectaron un número alto de ingresos por neumonía atípica, en un periodo corto de tiempo. Las investigaciones posteriores llevaron a la definición de una nueva entidad, el Síndrome del Aceite Tóxico Español. Con frecuencia el problema *aparenta* estar bien definido. Por ejemplo, en un estudio sobre factores dietéticos y metabólicos y su influencia sobre la litiasis renal, será necesario diferenciar el tipo de litiasis: no se va a detectar lo mismo si se trata de ácido úrico que de oxalato cálcico.

Si el conocimiento existente sobre el problema identificado es escaso, el estudio más indicado será de tipo descriptivo. Se trata de medir el problema buscando diferencias entre los distintos grupos y subgrupos, en función de la desigual distribución de otras variables. Los estudios que consideraremos en este caso son los **Ecológicos** y **Transversales**.

Cuando existen suficientes evidencias a partir de diseños descriptivos, se está en condiciones de dar un paso más. Se trata ya de evaluar la posible asociación, de tipo causal, entre diferentes factores y el problema de salud de interés. ¿Qué factores? Probablemente la etapa descriptiva ha sugerido ya un buen número de candidatos, aunque no es la única fuente de hipótesis. En numerosos casos la observación clínica atenta ha proporcionado pistas valiosas para la identificación de factores de riesgo y factores pronósticos. En este estadio de la investigación habrá que optar entre dos grandes grupos de estudios: los **Experimentales** y los **No experimentales**. En el primer caso están incluidos los **estudios aleatorios controlados** (también conocidos como ensayos clínicos) y los diseños **cuasiexperimentales**. En el grupo de los no experimentales (también denominados, por algunos autores, *Analíticos*) se encuentran los estudios de **seguimiento** y los de **casos y controles**. La elección entre uno u otro grupo dependerá de los objetivos, de criterios éticos, de los recursos y el tiempo disponibles, y otros factores que se comentarán más adelante.

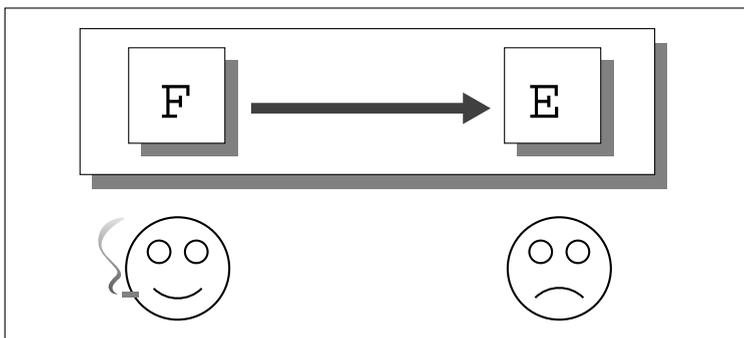


Fig. I.3- Asociación entre un factor de exposición (F) y un efecto (E).

El objetivo a esta altura de la investigación es la verificación de hipótesis. Se intenta averiguar si hay una determinada asociación (Fig. I.3) y si existen evidencias de que esa relación es de tipo causal y puede por tanto, hablarse de factores pronósticos o factores de riesgo.

3. Tipos de estudios: La Observación y la Experimentación

En la secuencia que parte de la definición del problema a investigar se recorren los principales tipos de estudios en Epidemiología. Sin poder hacer una presentación exhaustiva, se intentará ahora sistematizar un esquema general de las opciones que más frecuentemente se presentan al investigador³ (Fig. I.4).

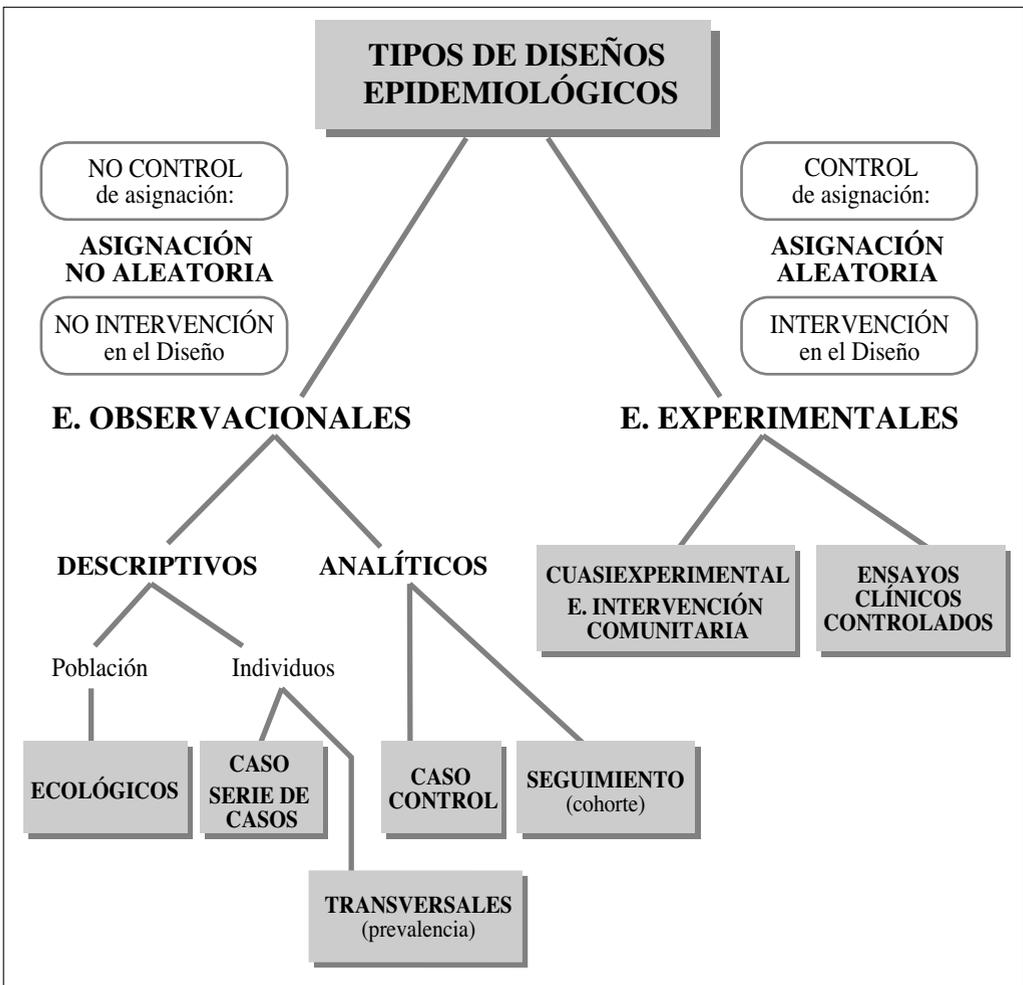


Fig. I.4- Tipos de estudios. El control de la asignación y la intervención.

Una primera gran división es la que distingue entre diseños **Experimentales** y los **No Experimentales** u **Observacionales**. Se utilizan dos criterios para establecer esta distinción. En primer lugar, la posibilidad de que el investigador controle o no la asignación de los participantes a los grupos de comparación. En un estudio experimental existe ese control, es decir, el propio investigador decide, para los individuos que van a participar, a qué grupo serán asignados. Aunque, como se verá más adelante, el control en la asignación se deja en manos del azar (lo que conocemos como *asignación aleatoria*). En los diseños no experimentales no hay posibilidad de asignación aleatoria; las personas presentan exposición o no a distintos factores, por las razones que sean, sin que intervengan los investigadores. Estos se limitan a observar la situación.

Por otro lado, en los estudios de tipo experimental existe manipulación de alguna variable independiente, o **intervención**, por parte del investigador. Es decir, se espera que tras la asignación aleatoria, las diferencias que se encuentren se puedan atribuir, precisamente, a la intervención efectuada.

4. Estudios experimentales

Si hubiera que elegir cuál es el mejor diseño posible cuando se pretende verificar una hipótesis, la respuesta sería, en principio, favorable a los diseños de tipo experimental. La razón reside en sus características. En primer lugar, el **experimento** permite el control de la asignación. El hecho de poder incluir a las personas en los grupos que se van a comparar de forma aleatoria tiene una ventaja sobre el resto de las estrategias: si el azar no juega una mala pasada, se puede esperar que los grupos de comparación sean muy similares entre sí. Es decir, para múltiples variables (como sexo, edad y demás) es posible conseguir grupos homogéneos, en los que la distribución de las variables que podrían interferir en los resultados sea semejante. De esta forma, si se encuentran diferencias, sólo quedaría una explicación: la intervención, o manipulación de una variable. La situación ideal la representa, pues, el **estudio aleatorio controlado**^{4,5} (Fig. I.5).

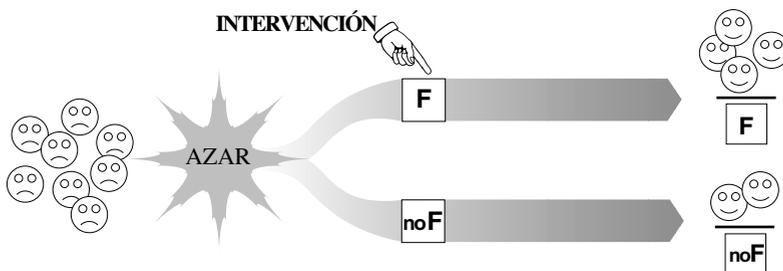


Fig. I.5- Estudio aleatorio controlado (ensayo clínico). Tras la asignación aleatoria, la exposición (F) consiste en una intervención. Al final se comparan los resultados obtenidos en términos relativos.

Una opción intermedia es la que ofrecen los estudios **cuasiexperimentales** o **casierperimentales**⁶. En ellos puede no existir asignación aleatoria individual; como mucho, puede decidirse al azar cuál será el grupo en el que se intervendrá. Pero una vez decidido ésto, todas las personas que integran cada grupo recibirán o no la intervención. Es el caso en el que se pretendiera evaluar el impacto de una intervención preventiva sobre la calidad de vida de un tipo de pacientes. En el experimento puro, se distribuirían individualmente al azar en dos grupos, uno de ellos sometido al nuevo tipo de intervención, y el otro a la intervención estándar preexistente. En el cuasiexperimento, los enfermos tratados en distintos servicios, o residentes en varias zonas geográficas, recibirían la intervención sólo si en su servicio o su zona se decidió que se intervendría. Incluso es posible el diseño cuasiexperimental en el que no exista grupo control. En ese caso sólo queda la posibilidad de evaluar la situación *antes* y *después* de la intervención.

Aún cuando los estudios experimentales pueden parecer enormemente atractivos por sus ventajas, en multitud de ocasiones son impracticables. Por un lado, razones éticas impiden exponer a seres humanos a situaciones que pudieran ser perjudiciales para ellos. Por ello, si la hipótesis a evaluar plantea que la presencia de cierto factor incrementa el riesgo de que se desarrolle una enfermedad, o se piensa que cierta terapéutica podría estar implicada en un peor pronóstico, cualquier tipo de experimentación está expresamente prohibida. Por otro lado, cada vez es más frecuente que la investigación etiológica se dirija hacia exposiciones ligadas a formas y estilos de vida. En ese caso, aún cuando se pensara que cierta modificación pudiera ser beneficiosa, resultaría prácticamente imposible conseguir que individuos asignados aleatoriamente modificaran sus hábitos y costumbres.

5. Estudios no experimentales

Estas y otras razones han hecho que muchas investigaciones utilicen estudios **No Experimentales** u **Observacionales**. En las últimas décadas se ha producido un aumento en el desarrollo de nuevos diseños, y un incremento espectacular de procedimientos de análisis estadístico para el control de variables ‘distorsionadoras’ o confundentes. No se debe olvidar, por otro lado, que el objetivo de la investigación no es necesariamente la verificación de hipótesis; de hecho persiste un gran desconocimiento sobre el comportamiento de los problemas que afectan a la Salud en las poblaciones, lo cual justifica el uso de estudios descriptivos.

Los diseños observacionales se subdividen a su vez en dos grandes tipos: los meramente **Descriptivos**, y los denominados **Analíticos**. En el primer caso se pretende encontrar la respuesta a una pregunta ya conocida: ¿cómo se comporta el problema

de interés en un determinado grupo? Como se ha comentado, la descripción de la situación en términos de características de persona, lugar y tiempo permite hacerse una idea de cuál es la situación. Esta aproximación descriptiva aporta un conocimiento sobre cuáles son los problemas y necesidades de atención existentes, además de la posibilidad de formular hipótesis que otros estudios, de tipo analítico, puedan evaluar.

En el grupo de los estudios descriptivos se plantean estrategias diferenciadas. Por un lado es posible manejar datos agregados, generalmente ya existentes cuando se inicia la investigación. Es el caso de los estudios **Ecológicos**⁷. En ellos no se usa información individualizada, sino que la unidad de análisis es el grupo o la población, y la información sobre la distribución de la exposición es el promedio dentro de cada unidad de estudio. Por ejemplo, un estudio podría comparar las ventas de tabaco en las distintas provincias españolas con la incidencia o la mortalidad por cáncer de vejiga en cada una de ellas. Se esperaría encontrar una relación entre mayor venta y, por tanto, consumo per cápita y una mayor presencia de ese tipo de cáncer. Los datos procederían de cualquier anuario estadístico, y no sería posible usar información desagregada por individuos. Eso diferencia a los estudios de tipo ecológico del resto de las aproximaciones descriptivas, en las que sí se dispondrá de datos individuales.

Otra opción descriptiva es partir directamente de las observaciones clínicas. Es frecuente la aparición de casos aislados *raros*, o de grupos de enfermos con un mismo diagnóstico y algún común denominador no observado previamente. Es lo que conocemos como **Informes de casos** y **Series de casos**. A pesar de su abundancia en la literatura, representan la opción más débil, con una utilidad limitada.

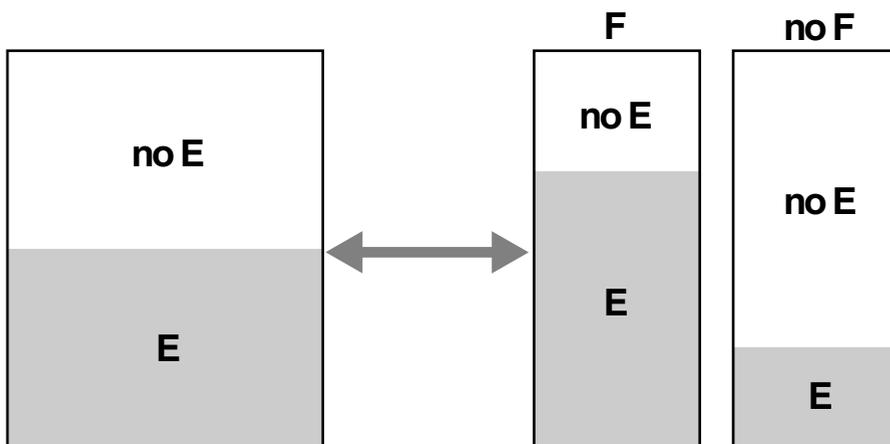


Fig. 1.6- Estudio transversal. La prevalencia de la enfermedad (E) puede compararse en expuestos (F) y no expuestos (no F).

Finalmente, se puede intentar el planteamiento de los estudios **Transversales** o de **Prevalencia**⁸ (Fig. I.6). Usando también información individualizada, se trata de conocer para una población definida (sean enfermos o sanos), qué proporción de individuos presentan una determinada característica o problema en un momento dado en el tiempo. La existencia o no del problema (*E* o *no E*) se contrasta simultáneamente con la presencia o no de una o más características (*F* o *no F*). Un estudio transversal sobre prevalencia de hipertrofia benigna de próstata podría encontrar, por ejemplo, una mayor proporción de personas con niveles altos de andrógenos (*F*) entre los individuos con hipertrofia (*E*), en comparación con los que presentan una próstata normal (*no E*).

Los estudios descriptivos dan lugar generalmente a hipótesis variadas. ¿Cómo verificarlas? Ya se han comentado las limitaciones y dificultades que se encuentran cuando se piensa en los estudios experimentales. En este punto, el investigador tiene la posibilidad de acudir a los diseños **Analíticos**. Incluidos en el grupo de los estudios Observacionales, y por lo tanto no experimentales, reúnen un conjunto de alternativas suficientemente potentes frente a la experimentación. Los más frecuentemente usados son los estudios de **Seguimiento** o de **Cohortes**, y los estudios de **Casos y Controles**.

En el caso de los estudios de **Seguimiento**⁹, el planteamiento es simple (Fig. I.7). Se parte de una o varias hipótesis que predicen la asociación entre una exposición (*F*) y algún efecto (*E*). Si esa hipótesis es cierta, es razonable esperar que las personas que presentan tal exposición tengan una mayor probabilidad de desarrollar el efecto después de un periodo más o menos largo. Podría compararse su experiencia con la de otro grupo en el que no estuviera presente la exposición (*no F*). Se dispondría entonces de, por lo menos, dos grupos: uno expuesto y el otro no; esos grupos se denominan *cohortes*.

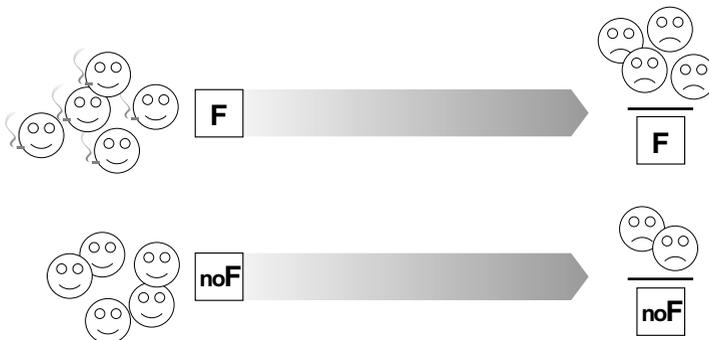


Fig. I.7- Estudio de seguimiento o de cohortes. Tras el seguimiento se compara la frecuencia de nuevos casos en expuestos (*F*) y no expuestos (*no F*); el denominador es la gente que había al comienzo libre de enfermedad.

Ya sólo queda seguir a ambas cohortes en el tiempo y verificar si efectivamente, en términos relativos, se produce un mayor número de nuevos casos en la cohorte expuesta en comparación con la no expuesta. Esta forma de ver las cosas difiere muy poco de lo que se hace en un estudio aleatorio controlado. Sólo un detalle: aquí los investigadores no deciden quién se expone y quién no. Por distintas razones algunos individuos pertenecen a una cohorte o a otra, y sólo es posible *observar* qué ocurre a lo largo del tiempo. Como consecuencia, tampoco existe otra *intervención* que no sea la derivada de la observación y de la medición.

Los estudios de seguimiento, como más adelante se comentará, plantean dificultades diversas. Entre ellas, el amplio tiempo de seguimiento que se requiere especialmente cuando el análisis se centra en enfermedades con largos períodos de inducción (desde que se inicia la exposición hasta que aparece el efecto). Por ello, los diseños del tipo **caso-control**^{10, 11} son una opción frecuentemente preferida (Fig. I.8).

La estrategia en un estudio de casos y controles, en su versión más clásica¹², se fundamenta en lo siguiente: si es cierto que una exposición (F) está asociada con un aumento en la probabilidad de que aparezca cierto problema (E), es posible encontrar antecedentes de su presencia en las personas en las que el problema ha aparecido ya. Y esos antecedentes de exposición deberían ser mayores que en personas en las que el problema no ha aparecido (*no E*). La comparación se hace, en este caso, entre gente con o sin el efecto presente. Siguiendo con un ejemplo anterior, esperaríamos encontrar más antecedentes de exposición al tabaco (F) en los casos de cáncer de vejiga (E) que entre los controles (en este caso, las personas que no han desarrollado ese tipo de cáncer).

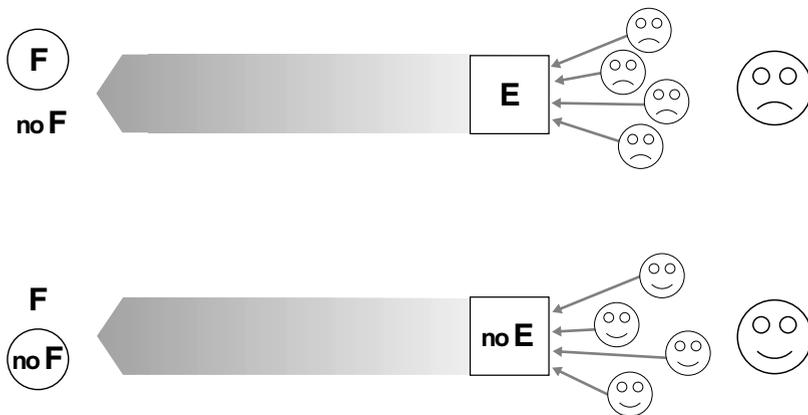


Fig. I.8- Estudio caso-control. Se comparan los antecedentes de exposición (F) y no exposición (*no F*) en casos (E) y en controles (*no E*).

Los diseños observacionales de tipo analítico son, pues, una buena opción cuando se intenta evaluar hipótesis. Aún cuando no pueden aprovechar las ventajas de la asignación aleatoria y el control de las intervenciones, serán el diseño de elección en múltiples ocasiones. El problema va a ser, ya que no existe experimentación, en qué medida otros factores pudieran estar distorsionando la observación.

B. ERRORES EN LOS DISEÑOS Y FORMAS DE CONTROL

E. Perea-Milla López

El objetivo de la mayoría de los diseños epidemiológicos es aprender algo acerca del efecto de una exposición sobre un problema de salud ¹³. En general, el objetivo final del proceso de investigación es encontrar diferencias en los grupos que se comparan para poder decidir si existe una asociación, y si esa asociación es de tipo causal. En la medida en que las diferencias observadas en un estudio difieran de las diferencias reales, el estudio contiene algún error.

Los posibles errores se agrupan en dos grandes tipos: error **aleatorio** y error **sistemático** ¹⁴.

1. Error aleatorio

El *error aleatorio* afecta a la precisión de las estimaciones y se produce cuando el estudio se realiza sobre una muestra probabilística de la población. Si se trabaja con muestras, es difícil esperar que los resultados *coincidan exactamente* con lo que ocurre en la población. De forma que lo obtenido en la muestra puede ser algo mayor o algo menor que lo que se hubiera observado de haber trabajado con toda la población. Con datos procedentes de una muestra nunca se puede saber si el error aleatorio ha actuado en un sentido u otro, sólo se sabe que existe.

La estadística aporta instrumentos para controlar la precisión de cualquier estudio. En una primera fase, mediante el cálculo del mínimo *tamaño de muestra* necesario para poder detectar como estadísticamente significativa una diferencia si es que ésta realmente existe en la población. En segundo lugar, mediante el uso de *pruebas o tests de hipótesis* y el cálculo de *intervalos de confianza* para las estimaciones obtenidas (Capítulo II.C). Con esta forma de trabajar se consigue que los diseños tengan la mayor precisión posible, es decir, que las asociaciones que realmente existen en la población estudiada aparezcan como estadísticamente significativas y los intervalos de confianza sean lo más estrechos posible.

2. Error sistemático o sesgo

El *error sistemático* desplaza artificialmente las diferencias observadas en el estudio de las verdaderas, se muestree o no. Ese desplazamiento se produce siempre en alguna dirección. En unas ocasiones, en el sentido de *aumentarlas*; cuando éso ocurre, el estudio exagera las verdaderas diferencias e incluso detecta algunas que en realidad no existen. Otras veces, tiende a *minimizarlas*, y entonces el estudio detecta diferencias menores a las reales o no encuentra las que realmente existen. En cualquier caso, la presencia de errores sistemáticos afecta a la **validez interna** del diseño, es decir, la credibilidad de las conclusiones respecto a los individuos estudiados. Con frecuencia pueden coexistir varios errores sistemáticos en un mismo estudio.

Se han descrito multitud de errores sistemáticos o **sesgos**¹⁵. Se acepta en general que pueden agruparse en tres categorías: sesgos de **selección**, de **clasificación** y de **confusión**.

El **sesgo de selección** aparece cuando las personas incluidas en el estudio difieren en alguna característica relevante de la población a la que se pretende aplicar las conclusiones. Si esa característica está relacionada con el factor de exposición y con el efecto de interés, entonces los hallazgos no son totalmente extrapolables a esa población, y pueden además modificarse los resultados del estudio. Es el caso, por ejemplo, de las negativas a participar, o los abandonos durante un seguimiento.

La **mala clasificación** es producto de cualquier ejercicio de medición y puede afectar tanto a la **exposición** como al **efecto**. Este sesgo, conocido también como sesgo de *información*, depende de la validez y fiabilidad del método utilizado para recogerla. Su impacto puede ir en el sentido de diluir diferencias realmente existentes, si se produce por igual en todos los grupos de comparación (con algunas excepciones). En ese caso se conoce como sesgo de mala clasificación **no diferencial**. Si la gente tiende a ocultar o exagerar hábitos socialmente no admitidos, es posible, por ejemplo, que un estudio no detecte de forma clara quiénes están expuestos o no a prácticas de riesgo para las enfermedades de transmisión sexual. Al mezclar erróneamente gente con o sin la exposición, es más difícil encontrar asociaciones si es que existen.

Si, por el contrario, el error es más frecuente en alguno de los grupos en estudio, es decir, está condicionado por la presencia o ausencia del efecto o la exposición, entonces tiende a sobreestimar o infraestimar las diferencias. Se trata entonces de una mala clasificación **diferencial**. Un caso clásico en la literatura es el de las madres de niños nacidos con malformaciones. Probablemente, recuerden con mayor intensidad las exposiciones ocurridas durante el embarazo que las que tuvieron hijos sanos, produciéndose en este caso una sobreestimación del verdadero riesgo asociado.

El tercer tipo de sesgo es el de **confusión**. Algunos autores lo consideran distinguiéndolo de los anteriores, ya que es el único que se puede controlar en la fase de análisis y no sólo en la de diseño. Se trata de una distorsión en las estimaciones del estudio, producida por la distribución desigual, en los grupos de comparación, de otra variable (tercera variable o variable *confundente*). Si esa variable es a su vez predictora del efecto (factor o marcador de riesgo, o protector) entonces, su distribución no balanceada (desigual) contamina la verdadera relación entre la exposición y el efecto que se están evaluando.

Un ejemplo puede encontrarse en el cáncer de vejiga (Fig. I.9). Se piensa que entre el 30 y el 70% de los casos están relacionados con el consumo de tabaco¹⁶ y un 10 o un 20%, con ciertas exposiciones laborales (industrias textiles, mecánicas y gráficas sobre todo). Se sospecha además que algunos factores dietéticos pueden estar implicados.

Se podría evaluar mediante un estudio la asociación entre cierto tipo de dieta (digamos, una dieta *X*) y ese cáncer. El diseño debería tener en cuenta la posibilidad de que variables confundentes distorsionen sus resultados. Puede ser que entre quienes consumen ese tipo de dieta (*F*) sea más frecuente encontrar a trabajadores de las industrias en las que se sabe que existe un mayor riesgo de padecer cáncer de vejiga. Si éso pasa inadvertido para los investigadores, entonces parte de la *carga* del riesgo laboral se camuflaría junto a la de la dieta, aumentando aparentemente la peligrosidad de esta última.

Una distorsión semejante, pero de signo contrario, ocurriría si quienes siguen esa dieta son, por ejemplo, menos fumadores en promedio. Cuando se compare la dieta *X* con otras, la menor cantidad de tabaquismo en ese grupo provocará que esa dieta aparezca como menos peligrosa de lo que realmente es. Esto es así porque al asociarse la dieta en estudio con una menor proporción de fumadores, parte de la carga de riesgo del tabaco se transfiere en el estudio a las personas que siguen otros tipos de dieta.

Como puede verse, la distorsión que introduce un factor de confusión puede ir en el sentido de *exagerar* diferencias, existan o no en la realidad. Pero también puede actuar en el sentido contrario, y disminuirlas o *diluir*las. Sin embargo, en cualquiera de esas situaciones, la confusión consiste en una desigual distribución, en los grupos que se comparan, de algún factor que, además, se relaciona con el efecto estudiado.

¿Cómo decidir, en la fase de diseño, si una variable puede actuar como confundente?. La identificación o caracterización de una tercera variable como potencialmente confundente (*C*), es decir, distorsionadora de la verdadera relación entre el factor en estudio (*F*) y el efecto (*E*), requiere el cumplimiento de una serie de condiciones.

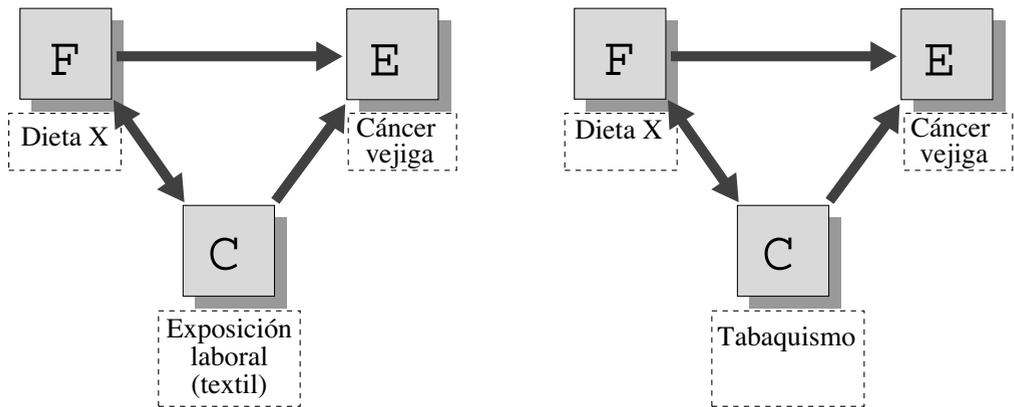


Fig. I.9: Factores confundentes (C). Su relación con la exposición (F) y el efecto (E) en estudio.

En primer lugar, el potencial confundente (C) debe ser predictor del efecto aún entre las personas no expuestas al factor en estudio. En el ejemplo anterior (Fig. I.10), tanto el tabaco como el trabajo en ciertas industrias se comportan, según estudios previos, como factores de riesgo (por tanto, predictores) en relación a la probabilidad de desarrollar cáncer de vejiga, aún entre aquéllos que no sigan la dieta X. Un factor confundente es aquél que actúa como factor de riesgo o factor protector del efecto en estudio, por sí mismo o como marcador indirecto de otro predictor.

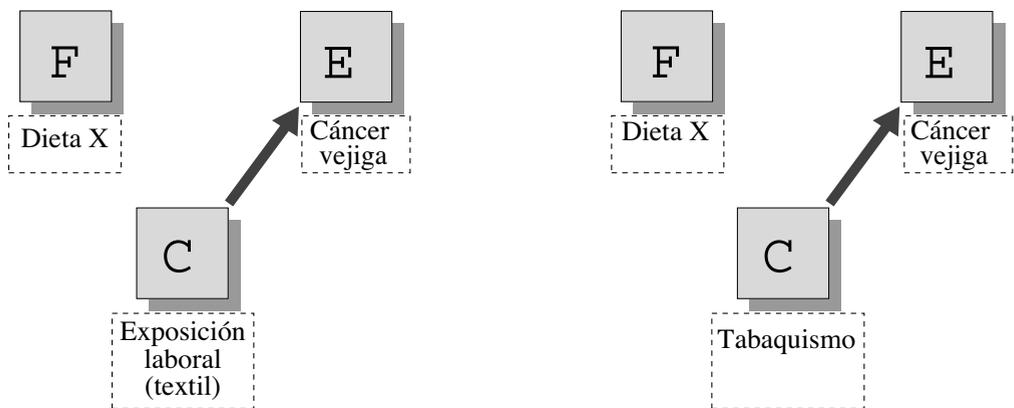


Fig.I.10: Factores confundentes y su relación con el efecto.

En segundo lugar, la variable confundente debe estar asociada con la exposición en estudio aún entre quienes no desarrollan el efecto. Es decir, debe sospecharse que entre las personas que siguen esa dieta es común encontrar a trabajadores de esas industrias, o que es normal encontrar menos gente que fuma (Fig. I.11), independientemente de que desarrollen o no cáncer de vejiga. Esto puede deberse a mayor o menor nivel cultural, diferentes estilos de vida, o cualquier otra razón.

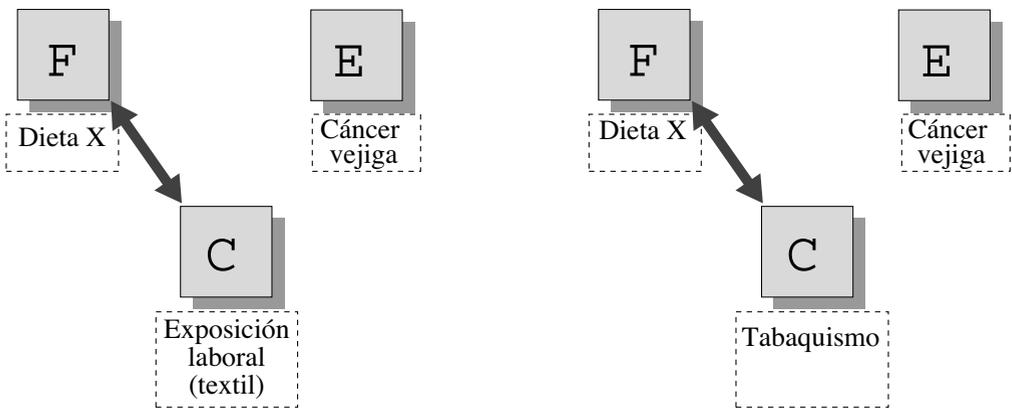


Fig. I.11: Factores confundentes y su relación con la exposición.

Rothman¹⁴ sugirió una tercera condición para evaluar el carácter confundente de una variable: que no supusiera un eslabón intermedio en la ruta causal entre la exposición (F) y el efecto (E). Esto es difícil de establecer en ocasiones, y el investigador sólo puede contar aquí con un buen marco teórico en el que fundamentar su argumentación. La cuestión es más complicada cuando la variable potencialmente confundente pudiera jugar un papel parcial en la hipotética ruta causal¹⁷.

3. Control de los errores

Los sesgos de selección y de clasificación deben controlarse en la medida de lo posible durante la fase de diseño del estudio, de lo contrario, pueden llegar a invalidar los resultados.

También la confusión puede controlarse en la fase de diseño. Se utiliza en ocasiones la *restricción*, es decir, la limitación del estudio a un sólo subgrupo. Si el sexo fuera un confundente ya conocido, se puede centrar el estudio en hombres o mujeres únicamente. Otra posibilidad de control es mediante lo que se conoce como *aparea-*

miento, haciendo que los grupos de comparación sean similares en cuanto a factores confundentes ya conocidos. Por ejemplo, elegir cada control con una edad semejante al caso correspondiente. La mejor herramienta en el diseño es, sin embargo, la asignación aleatoria, algo que no siempre es posible.

Si existió la suficiente precaución en la fase de diseño, y se recogió información sobre las variables potencialmente confundentes, es posible controlar entonces la confusión en el análisis de los datos usando técnicas estadísticas. Entre éstas se encuentran el *análisis estratificado*, el *ajuste de tasas* y el *análisis multivariante*.

En las próximas secciones se tratarán los sesgos de selección y de clasificación más frecuentes en los distintos tipos de estudio. Del control estadístico de la confusión se ocupa el Capítulo II.E.

La precisión del estudio y su validez interna son indicadores de la calidad de cualquier diseño. Determinan además el grado en el que sus conclusiones pueden extrapolarse a otras poblaciones distintas a la estudiada. Eso es lo que se conoce como **validez externa**. Estudios poco precisos y/o con poca validez interna no tienen mayor interés en cuanto a la validez externa. Ésta requiere el uso de razonamientos acerca del grado en el que los resultados pueden aplicarse a otras poblaciones a partir de lo obtenido en un estudio suficientemente preciso y válido.

C. MEDIDAS DE FRECUENCIA

E. Perea-Milla López

1. Introducción

Este capítulo presenta las herramientas básicas de medición empleadas en la descripción y comparación de grupos. El lector poco interesado en profundizar en el tema puede entender el resto de los capítulos sin detenerse aquí, aunque su lectura es aconsejable para la interpretación precisa del contenido de otras partes de la monografía.

En el manejo de los distintos tipos de estudio, es clara la necesidad de un instrumento que permita la medición y comparación de grupos. Las medidas de frecuencia cumplen esa función. Tales medidas caracterizan la ocurrencia de sucesos como enfermedad, incapacidad, muerte o cualquier otro problema relacionado con la salud, y son de utilidad en investigaciones experimentales y no experimentales (u observacionales).

El primer problema que se plantea al usar medidas de frecuencia para comparar grupos es que la utilización de números absolutos (por ejemplo, ocurrencia de un número dado de sucesos) es en general inadecuada. Se requiere que esas medidas sean independientes del tamaño de la población, lo cual implica que deben utilizarse junto a **denominadores**, es decir: el número de sucesos o fenómenos ocurridos debe dividirse por el número de individuos que componen las poblaciones estudiadas. Esos cocientes, las medidas de frecuencia relativa, serán el instrumento básico cuando la comparación de grupos persiga evaluar hipótesis y, en general, en cualquier comparación de grupos. Es importante tener claro que el concepto de población no necesariamente se refiere a población general en sentido demográfico. Cualquier conjunto de individuos definidos con algún criterio es una población (por ejemplo, pacientes que acuden a una consulta en un periodo de tiempo determinado).

Las distintas medidas de frecuencia tienen diferentes características según qué se quiera medir. Así, es posible medir la situación en un punto conceptual en el tiempo: tal medida de frecuencia se conoce como **Prevalencia**, y describe en qué proporción de la población estudiada está presente el suceso de interés, en un punto específico en el tiempo. Un buen símil para la prevalencia la compara con una «fotografía» de esa población (Fig. I.12). Por ejemplo, se estima que en EEUU la prevalencia de la

enfermedad de Peyronie es de 388,8 por 100.000 habitantes. Es decir, de cada 100.000 habitantes, se espera encontrar a casi 400 de ellos con esa enfermedad.

Otra posibilidad es medir lo que está pasando durante un periodo de tiempo: a esta medida se le denomina **Incidencia**; describe la frecuencia de nuevos casos que aparecen durante un periodo de tiempo. De alguna forma, la incidencia informa sobre el flujo que se produce en el interior de una población desde una situación (por ejemplo salud) a otra situación (en el mismo ejemplo, enfermedad). El símil en este caso es el de una «película», de principio a fin. En el caso de la enfermedad de Peyronie, aparecen unos 32.000 nuevos casos cada año en EE.UU. Lógicamente, para comparar lo que ocurre en EE.UU y España, sería necesario tener en cuenta el tamaño de la población masculina de ambos países.

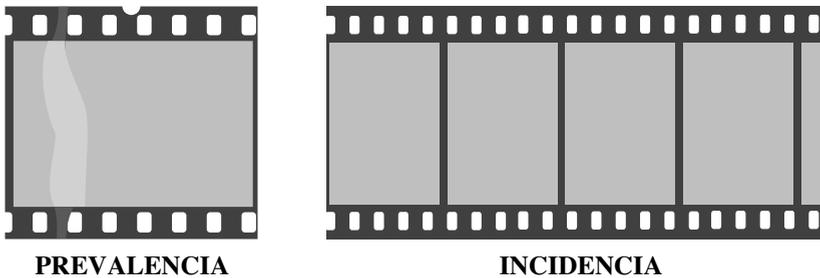


Fig. I.12- Prevalencia (Fotografía) e Incidencia (Película).

Aún siendo diferentes tipos de medidas, existe una evidente interrelación entre la incidencia y la prevalencia (Fig. I.13). Intuitivamente, es claro que la prevalencia de un problema de salud dado depende, por un lado, de la mayor o menor incidencia (aparición de nuevos casos) en esa población y, por otro lado, de su duración promedio (tiempo durante el que los individuos permanecen afectados hasta que el problema desaparece por curación, muerte,...).



Fig. I.13- Relacion entre incidencia, prevalencia y duración promedio de la enfermedad.

Es importante conservar en mente la distinción entre prevalencia e incidencia. Entre otras razones, porque sirven a propósitos diferentes. Los datos sobre **prevalencia** son más útiles desde el punto de vista del sistema de salud para conocer el impacto de los distintos problemas que afectan a la salud y poder predecir necesidades de atención, asignar recursos, priorizar intervenciones, etc. Además, las enfermedades consideradas «crónicas» (en general, las no transmisibles) ven mejor reflejado su impacto sobre la población si se maneja información sobre su prevalencia.

En el otro extremo, la incidencia permite una mejor aproximación a la magnitud del daño a la salud que representan las enfermedades consideradas como «agudas». Asimismo, está más indicado manejar incidencia en investigación etiológica y evaluativa. La razón es que el intentar verificar hipótesis usando casos prevalentes (personas afectadas por el problema, antiguas y nuevas) puede llevar a conclusiones erróneas. Si la exposición al factor de riesgo que se intenta valorar está relacionada con la duración del efecto (por ejemplo, si una cierta exposición mejorara el pronóstico de los casos), aparecería entonces más frecuentemente en el grupo de las personas afectadas. Este argumento es igualmente válido cuando se persigue evaluar el efecto de medidas preventivas.

Una vez conocido el interés que un tipo y otro de medidas tiene, es necesario explorar la estructura interna y las ventajas e inconvenientes que cada una de ellas aporta.

2. Prevalencia

Tal como se ha definido, consiste en el cociente:

$$P(E) = \frac{\text{Número de Personas que presentan el problema en un momento en el tiempo}}{\text{Número de personas en la población en ese momento}}$$

Se trata de la **proporción** de gente estudiada que presenta el fenómeno de interés (E) en un cierto punto en el tiempo. Esa proporción $P(E)$, asimilable a **Porcentaje**, puede oscilar en un rango comprendido entre 0 y 1 (valores mínimo y máximo posibles) y, al estar presente la misma unidad de medida en el numerador y en el denominador, el cociente resultante es adimensional. Puede calcularse la prevalencia en un grupo, o comparar distintas prevalencias según la presencia o ausencia de determinados factores de exposición (Fig. I.14).

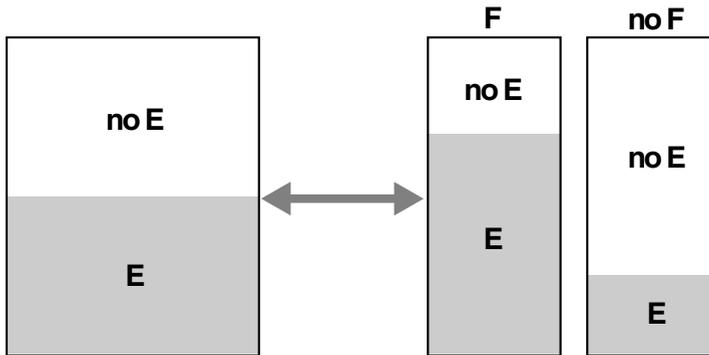


Fig. 1.14- Prevalencia: proporción de personas con el efecto (E) o sin él (no E) global, y según los grupos expuestos (F) o no (no F).

Su interpretación se traduce como la estimación de la **probabilidad promedio individual** de presentar el fenómeno. Por ejemplo, prevalencia de mujeres embarazadas de alto riesgo en una determinada comunidad a comienzos de año. Igualmente, probabilidad promedio individual de presentar hipertensión en el inicio del tercer trimestre de embarazo en una población de mujeres que no acuden a los servicios de control del embarazo.

Es de interés recalcar aquí una diferencia implícita en los dos ejemplos anteriores. El punto de corte conceptual no es estrictamente cronológico; es posible estudiar prevalencia de eventos que no se producen necesariamente en una fecha común para todas las personas estudiadas. Es el caso de la aparición de niños recién nacidos que presentan malformaciones renales. Si la presencia de ese tipo de malformaciones se asocia con algún grado de inviabilidad, es muy posible que un porcentaje de los embriones y fetos que a lo largo del embarazo desarrollen el problema no lleguen a alcanzar la viabilidad y se produzca muerte fetal intraútero, aborto... De tal forma que el número de recién nacidos vivos con esa malformación no es sino una expresión de la probabilidad promedio individual de presentar ese daño a la salud en el momento del nacimiento, es decir de su prevalencia. De la misma forma, un estudio que se limitara a describir la presencia de una determinada característica en el momento del diagnóstico en pacientes que acuden a una consulta, estaría midiendo la prevalencia de esa característica en esos pacientes.

3. Incidencia

Su cuantificación puede llevarse a cabo utilizando dos medidas distintas, dependiendo de qué información pueda recogerse. En cualquier caso, ambas tratan sobre la aparición de nuevos casos durante un periodo de tiempo.

3.1. Incidencia medida como INCIDENCIA ACUMULADA.

En este caso, la incidencia se define como:

$$IA = \frac{\text{Número de nuevos casos que aparecen durante un periodo de tiempo}}{\text{Número de individuos sin el problema al comienzo del periodo}}$$

Se puede definir como la **proporción** de personas en las que, no estando inicialmente presente, el problema de salud aparece durante un periodo de tiempo dado (Fig. I.15). Están incluidos sólo individuos que, al comienzo del periodo de seguimiento, estaban libres del problema y, por tanto, en riesgo de que éste apareciera. Es lo que se conoce como **población en estudio**.



Fig. I.15- Incidencia medida como incidencia acumulada en una población en estudio expuesta a un factor de riesgo (F).

Esta imposición, difícil de cumplir cuando se manejan datos poblacionales agregados y no se conoce con exactitud el número de personas inicialmente en riesgo, será una exigencia inevitable si lo que se plantea es un diseño específico de investigación, lo que se conoce como un estudio de seguimiento o de «cohortes». En ese caso, posiblemente sea necesario plantear un estudio de prevalencia al comienzo para eliminar del seguimiento a los ya afectados.

Como en el caso de la prevalencia, la incidencia, medida como incidencia acumulada, es una **proporción**. Y de la misma forma, puede oscilar en un rango comprendido entre 0 y 1, y es adimensional.

Se interpreta como el riesgo o **probabilidad promedio individual** de aparición del problema de salud, durante un tiempo que debe especificarse explícitamente. Deben resaltarse dos asunciones empleadas en este caso: primero, que se utiliza un grupo (o varios) que es **fijo**. Y segundo, que se requiere una referencia explícita a un periodo de seguimiento específico.

Siguiendo a Rothman¹⁴, pueden establecerse restricciones al uso de la incidencia acumulada, sobre todo cuando no se sigue a una cohorte fija (o cerrada) desde el momento en que ocurre la exposición. Aún siendo un tipo de medida válido e interpretable, la incidencia medida como incidencia acumulada plantea algunas limitaciones en su uso.

En primer lugar, será mayor o menor dependiendo de la longitud del periodo de seguimiento; es de suponer que, a mayor tiempo de seguimiento, un mayor número de nuevos casos puede ir apareciendo. Y el número de individuos que están siendo seguidos puede disminuir conforme pasa el tiempo, por muertes, cambio de residencia u otras causas.

Además, en seguimientos largos, el grupo de personas estudiado, tanto por envejecimiento como por cambio en otras variables, puede diferenciarse y dejar de ser representativo de la población de la que procede.

Más aún, en algunos estudios, ciertas personas estarán en riesgo de que aparezca el problema durante diferentes períodos de tiempo, y serán seguidas dentro de distintos períodos en el estudio; en el caso de una población en la que se produzcan salidas (por muerte, emigración,...) o entradas (nuevos nacimientos, inmigración,...), no es posible conseguir un denominador «fijo» adecuado. No se puede delimitar, por tanto, lo que hemos denominado *población en estudio* al comienzo del seguimiento.

Por último, es importante tomar en consideración el tiempo transcurrido hasta que aparece el efecto estudiado. Puede darse el caso de dos grupos que se están comparando durante un mismo periodo de tiempo; aún cuando en el punto final del seguimiento la incidencia de nuevos casos acumulada sea la misma, es posible que en uno de ellos la aparición del problema haya ocurrido de forma más rápida que en el otro. En multitud de enfermedades, este factor, la velocidad de aparición, es un asunto nada desdeñable. La figura I.16 presenta el caso hipotético de dos ciudades y la mortalidad por todas las causas. En ambas muere todo el mundo, pero cualquiera elegiría haber vivido en la ciudad B: allí la muerte tarda más en presentarse.

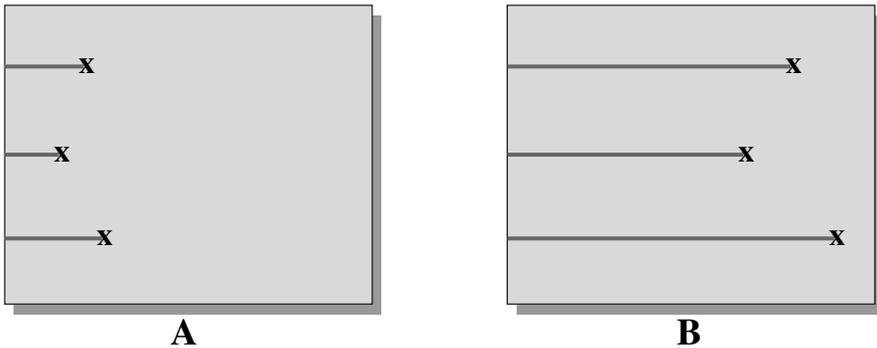


Fig. 1.16- Ciudad A y ciudad B. Mortalidad por todas las causas (X) durante el mismo periodo de tiempo¹⁴.

3.2 Incidencia medida como DENSIDAD DE INCIDENCIA.

El repaso de los problemas que puede plantear el manejo de la incidencia acumulada lleva a la constatación de que el número de personas estudiadas que pasan de la situación de «en riesgo» (ej.: sanos) a la de «suceso presente» (ej.: enfermos), durante cualquier periodo de tiempo, depende del tamaño de la población, de la longitud del tiempo de seguimiento, y de la mayor o menor velocidad en la aparición del cambio en cada población. Esta diferencia en la capacidad o «fuerza» para producir distintos flujos de cambio (de sanos a enfermos, en el ejemplo) dependerá de la presencia o ausencia, en distintos grados, de determinados factores de riesgo o exposición.

Es por esta razón que la **densidad de incidencia**, además de tener en cuenta la proporción de población que se afecta, toma en consideración el **tiempo** transcurrido hasta que aparece el efecto; en otras palabras, se incorpora la variable «tiempo en riesgo» a la estructura del cociente:

$$DI = \frac{\text{Número de nuevos casos que aparecen en una población durante un periodo de tiempo}}{\text{suma para cada individuo de la población del tiempo durante el que estuvo en riesgo de que apareciera el efecto}}$$

El numerador, ya que se está midiendo incidencia, es idéntico al que se utiliza en el cálculo de la incidencia acumulada, es decir, número de nuevos casos aparecidos.

Respecto al denominador, resulta obvio que cada persona seguida aporta al conjunto global todo el tiempo durante el que no presentó el suceso de interés, hasta que

éste apareció o bien, si ésto no ocurrió, hasta que el seguimiento terminó. El denominador es, entonces, una *suma de tiempo* (medido en años, días, meses, segundos,..), conocido como **personas-tiempo** (Fig. I.17).

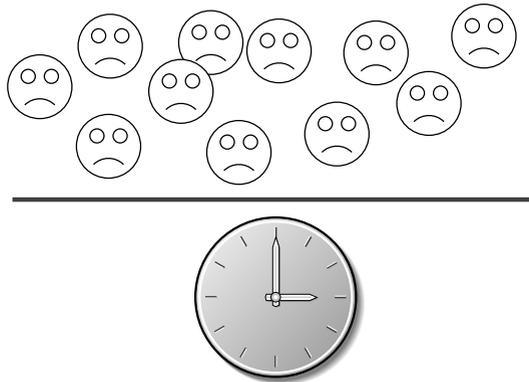


Fig. I.17- Densidad de incidencia. El denominador, tiempo en riesgo.

Es decir, se obtiene un sumatorio de los períodos de tiempo en riesgo observados en todos los individuos de esa población que está siendo seguida. Y esa suma será mayor o menor dependiendo tanto del tamaño de la población como de la duración del seguimiento.

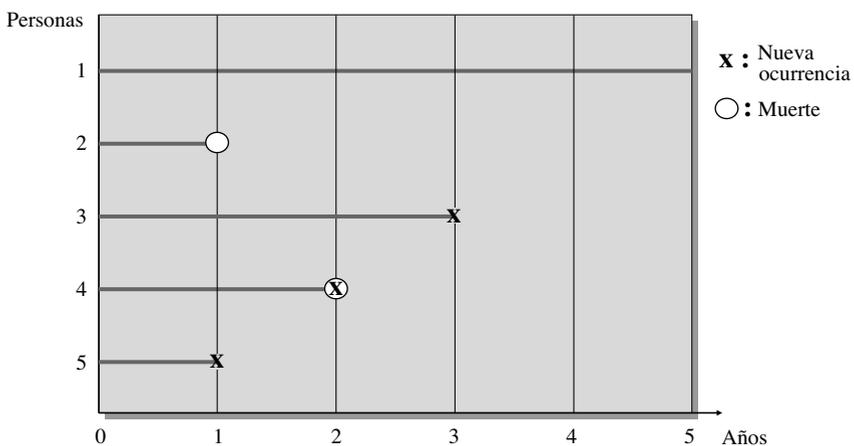


Fig. I.18- Densidad de incidencia. Cálculo del denominador (tiempo en riesgo).

La figura I.18 presenta un caso hipotético. Se sigue a 5 personas durante 5 años. Una de ellas (la número 1), llega al final del periodo sin padecer la enfermedad X. Tres desarrollan el efecto durante el seguimiento (números 3,4,5); la persona número

4 además fallece. La número 2 muere al final del primer año por otra causa. La cantidad de nuevos casos de enfermedad X aparecidos en esa población es de tres. Es el numerador. Para construir el denominador debe sumarse el tiempo que cada individuo permanece en seguimiento hasta que aparece la enfermedad, acaba el seguimiento, o desaparece por abandono, muerte por otra causa o cualquier otra razón.

El primer individuo aporta al conjunto de *tiempo en riesgo* 5 años. El segundo, sólo 1 año, pues muere. El tercero estuvo tres años en riesgo de enfermar, hasta que lo hizo: aporta 3 años. En total se han observado 12 *años en riesgo*. La densidad de incidencia es, pues, igual a $3/12=0.25$. O sea, la velocidad a la que este grupo ha enfermado (la velocidad en el cambio de sanos a enfermos) es de 0.25 por año de seguimiento, o año en riesgo. Haciendo una simple regla de tres, puede decirse que por cada 100 años de seguimiento que se consiga observar se producen 25 nuevos casos de enfermedad. No son estrictamente 100 años cronológicos; lo que se afirma es que de cada 100 años que se consigan reunir u observar se encontrarán 25 nuevos casos.

El resultado final de esta combinación informa, por lo tanto, sobre la mayor o menor **velocidad** de aparición del suceso de interés por unidad de tiempo de seguimiento, conocida como «fuerza de morbilidad» o «fuerza de mortalidad», según el caso. Se espera que ese potencial de cambio varíe de unas poblaciones a otras en función de los distintos grados de exposición presentes en cada una de ellas.

La figura I.19, tomada de Alvarez-Dardet et al¹⁸, ilustra la diferencia en velocidad en dos grupos distintos. Se supone dos grupos de enfermos en los que se prueban los tratamientos A y B.

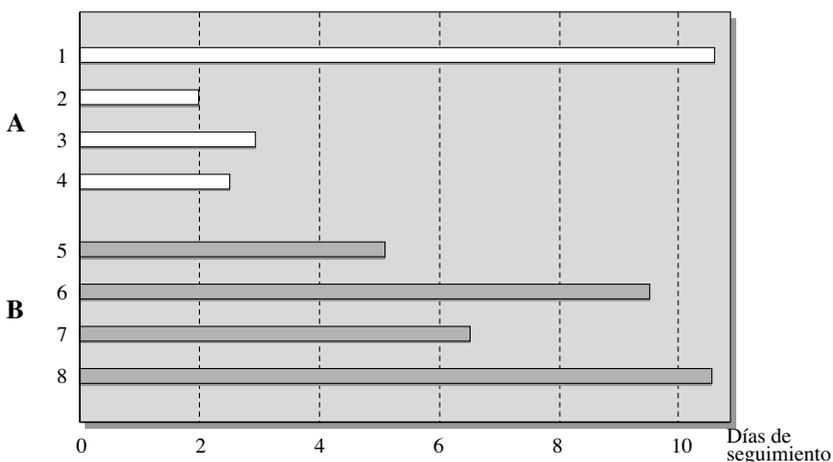


Fig. I.19- Densidad de incidencia. Diferencias en la velocidad de aparición del efecto (curación) según el tratamiento empleado.

En cada grupo se obtiene un 75% de curaciones después de 10 días de seguimiento. Sin embargo, el tratamiento A consigue que las curaciones se produzcan antes, aparezcan con mayor velocidad. El numerador para calcular la densidad de incidencia es 3 en ambos grupos. Sin embargo, el denominador (tiempo en *riesgo de curarse*) es de 15 días de observación en los enfermos sometidos a tratamiento A, y 30 días en el tratamiento B. De forma que la velocidad de curación es de $3/15 = 0.20$ por día en el primero y $3/30 = 0.10$ en el segundo. La velocidad de aparición de nuevos casos (en este ejemplo curaciones) es el doble si se usa el tratamiento A. Haciendo una regla de tres, se podría afirmar que por cada 100 días observados de tratamiento con A se obtienen 20 curaciones, frente a sólo 10 con el tratamiento B.

Otra forma de representar la diferencia de velocidad en aparición de nuevos casos es *alineando* los casos con su tiempo en riesgo correspondiente (Fig. I.20). En el grupo A puede comprobarse que la velocidad está siendo mayor que en el grupo B, que tiene una densidad de incidencia más baja. La aparición de nuevos casos es más lenta en B y por ello están más separados entre sí, ya que cada individuo aporta en ese grupo una mayor cantidad de tiempo en riesgo.

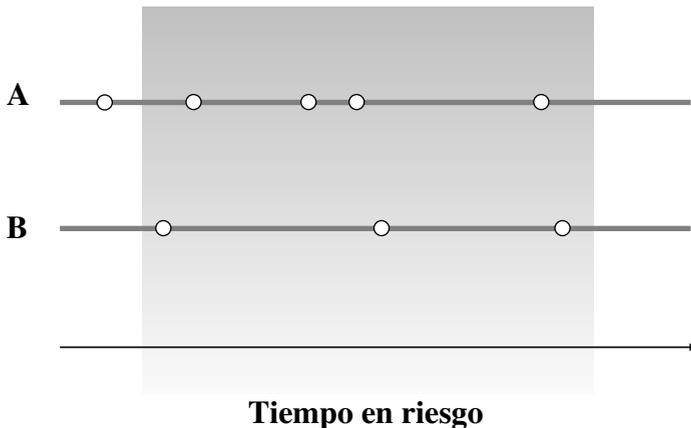


Fig. I.20- Tasa de incidencia. Densidad de puntos a lo largo del tiempo en riesgo.

La densidad de incidencia no cumple los criterios para ser considerada como **proporción**. Teóricamente, su rango oscila entre 0 e infinito, y además tiene dimensión: al ser frecuencia de aparición de nuevos casos por unidad de tiempo, su unidad de medida se expresa como **tiempo⁻¹**, es decir, «cambio por unidad de tiempo». Y no mide probabilidad promedio de aparición de nuevos casos, sino la fuerza de una exposición para cambiar en determinada población el flujo del estado de salud al de enfer-

medad; es una característica de esa enfermedad, que, como se ha visto, puede comportarse de forma diferente en distintos grupos que se comparen, según distintos niveles de exposición. Este tipo de cociente se conoce como **tasa**.

Esta restricción del término «tasa» a lo que se ha descrito como densidad de incidencia puede generar confusión ya que, históricamente, se ha utilizado asociada a cualquier medida de prevalencia o incidencia. La fuerza de la costumbre ha impuesto el uso indiscriminado de ese concepto que, hoy por hoy, se acepta.

Una ventaja añadida del uso de la densidad de incidencia es que es aplicable tanto en situaciones en las que se parte de una población fija (un número de personas caracterizadas porque inicialmente no presentan el efecto de interés), así como en el caso de las poblaciones llamadas *dinámicas* o *abiertas* (Fig. I.21).

En las poblaciones abiertas se admite que hay entradas y salidas a lo largo del seguimiento. Un buen ejemplo es el de una ciudad que se siguiera a lo largo del tiempo. La gente emigra, muere, nace, llegan inmigrantes, etc. O un hospital en el que a lo largo del tiempo se producen ingresos, estancias más o menos prolongadas, y altas o fallecimientos. Esa situación impide establecer un grupo inicial fijo, la *población en estudio*, por lo que un seguimiento *en tiempo real* no puede calcular incidencia acumulada. Sólo puede estimarse el denominador de *tiempo en riesgo* sumando las contribuciones individuales al conjunto de personas-tiempo. Ese sumatorio se define como la **base del estudio**.

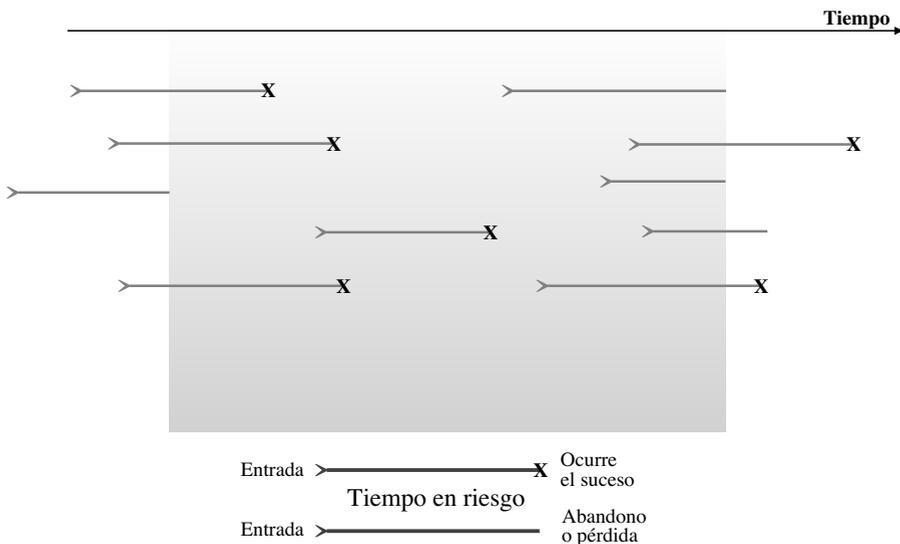


Fig. I.21- Población dinámica o abierta. Entradas y salidas durante el periodo de estudio.

Una población dinámica, con entradas y salidas, en la que ambas se contrapesan aproximadamente, puede considerarse como dinámica y estable. Si ésto es así, y puede asumirse que no hay cambios en la incidencia a lo largo del seguimiento, entonces cualquier periodo de observación que sea posible llevar a cabo, proporcionará una buena estimación de la incidencia del problema analizado en esa población.

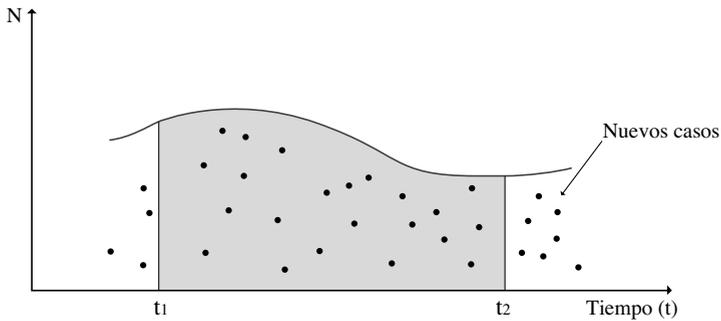


Fig. I.22- Densidad de puntos en el área de población y tiempo (la base de estudio).
En una población abierta, el tamaño cambia a lo largo del tiempo.

Tanto si la población es fija, como si es dinámica o abierta, la **base del estudio** o sumatorio de personas-tiempo puede representarse como el área definida por el tamaño de la población (*población de estudio*) en un eje, y el *tiempo de seguimiento* en el otro. En su interior se concentran, con mayor o menor *densidad*, los puntos que indican la aparición de nuevos casos, sea enfermedad, curación, recidivas o muertes. En una población abierta el tamaño del área oscila a lo largo del tiempo (Fig. I.22). Si es una población fija o cerrada, el tamaño es constante o decrece con el tiempo (Fig. I.23). Ese área es la *base de estudio*. Conceptualmente, podría estimarse la densidad de puntos en un periodo de tiempo (Δt) muy pequeño, por lo que la densidad de incidencia también se conoce como *tasa instantánea* (capítulo II.E.5).

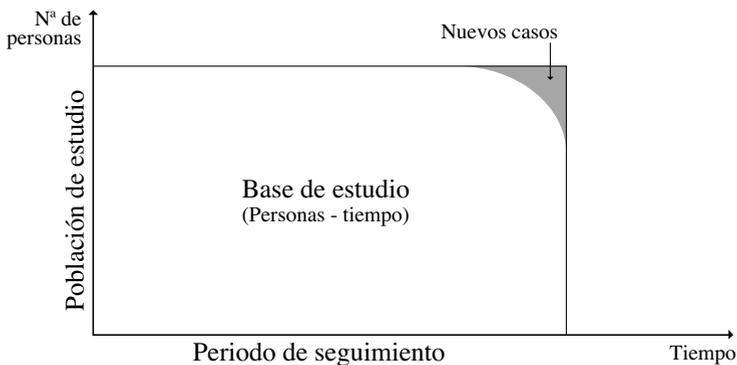


Fig. I.23- Población y base de estudio.
Fuente: Norell S⁴⁴.

En situaciones prácticas puede que no sea posible calcular el tiempo en riesgo para cada individuo. Una aproximación «grosera» a la construcción del denominador consiste en multiplicar el promedio de la población que se ha seguido, por el tiempo promedio de duración del seguimiento. Esa aproximación es frecuente cuando se usan datos agregados (a nivel poblacional) y no se dispone de información individual. Se basa en la suposición de que las entradas y salidas han sido homogéneas a lo largo del tiempo, y puede entonces asumirse que se han producido todas en la mitad del seguimiento (Fig. I.24).

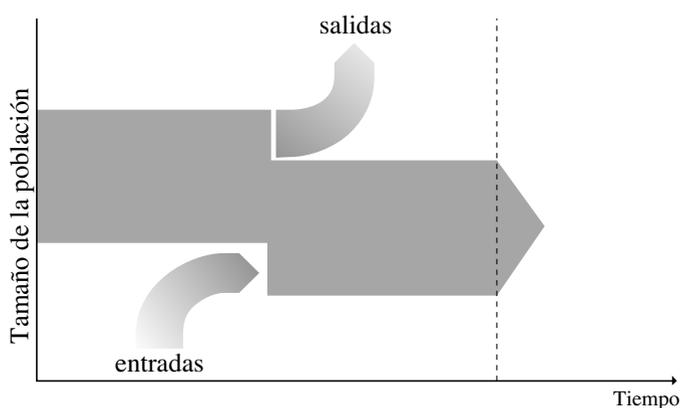


Fig. I.24- Población dinámica estable: entradas y salidas homogéneas durante el seguimiento.

4. Indicadores e índices

Las medidas de frecuencia revisadas previamente son medidas simples, obtenidas a partir de un sólo componente o variable. Parten de una sólo clase de datos, como podría ser la prevalencia de un determinado problema como la obesidad, o la incidencia de mortalidad en un cierto grupo poblacional. Las medidas así establecidas reciben el nombre de **Indicadores**. Tienen la ventaja de que resultan fáciles de relacionar directamente con los niveles de exposición a factores de riesgo o intervenciones realizadas por el sistema de salud. Aparentemente es fácil asociar cambios en hábitos de vida, como consumo excesivo de alcohol en embarazadas, con cambios en la proporción de recién nacidos de bajo peso.

Sin embargo, existe la posibilidad de construir medidas más complejas, denominadas **Índices**, que generalmente consisten en la combinación de distintos indicadores. Su objetivo es aproximarse de forma más precisa a la caracterización de una situación de salud determinada. Un ejemplo bien conocido de este tipo de medición

es el índice de **Quetelet**, que pretende determinar el mayor o menor grado de obesidad mediante el cociente entre el peso y la talla elevada al cuadrado. De esta forma, se ajusta el peso a la talla de cada individuo. Otro ejemplo de índice es el PSA densidad, que consiste en el cociente entre el nivel sérico de PSA y el peso en gramos de la próstata, medido mediante ecografía transrectal. Este otro tipo de combinación de medidas resume datos de dos o más componentes o variables. Es, por tanto, más complejo y multidimensional, al resumir componentes dispares.

Es posible construir múltiples tipos de índices. Cada uno de ellos podría combinar diferentes indicadores, y debería establecerse qué peso o ponderación se le atribuye a cada indicador. El problema práctico estriba, por lo tanto, en qué indicadores y qué pesos se decide usar, en función de lo que se pretende medir.

Cuando la medición se realiza en áreas geográficas o grupos poblacionales pequeños, sea con indicadores o con índices, aparece un problema adicional. Frecuentemente los datos desagregados no están disponibles o son inexactos, lo cual choca frontalmente con el interés en describir y planificar a ese nivel. En estas situaciones, por otra parte frecuentes, existe la opción de utilizar medidas de extrapolación. La idea es usar medidas de frecuencia (desde tasas a proporciones..) obtenidas de una población mayor (nacional, regional,..) para obtener una estimación sobre lo que se esperaría que ocurriera, por ejemplo, en un área local pequeña si ésta, con su propia estructura poblacional, hubiera tenido el mismo comportamiento respecto a morbilidad o mortalidad que el área tomada como referencia. Este método, conocido como **ajuste de tasas**⁸, se utiliza habitualmente tanto cuando el interés del estudio se centra en grupos pequeños, como cuando al comparar grupos entre sí existen variables (composición por sexo y edad) que se distribuyen en forma desigual y que interesa neutralizar.

D. ESTUDIOS DESCRIPTIVOS

E. Perea-Milla López

1. Introducción

Este tipo de diseños, que hemos situado en una perspectiva de primera aproximación al conocimiento de un determinado problema, se realizan por simple observación. Únicamente permiten, por lo tanto, constatar una situación dada. Consiguen sin embargo alcanzar dos tipos de objetivos. Por un lado, ofrecen información sobre la situación del problema de salud de interés, con la mayor precisión posible. De esta forma, identifican proporciones de individuos afectados y/o relacionan la distribución de algunas características o exposiciones con la del problema en estudio. Por otro lado, son una fuente inestimable de hipótesis, aún cuando su capacidad de verificación sea al menos discutible.

Se pueden sintetizar las distintas etapas en el diseño de cualquier estudio descriptivo, en un esquema simple (Fig. I.25).

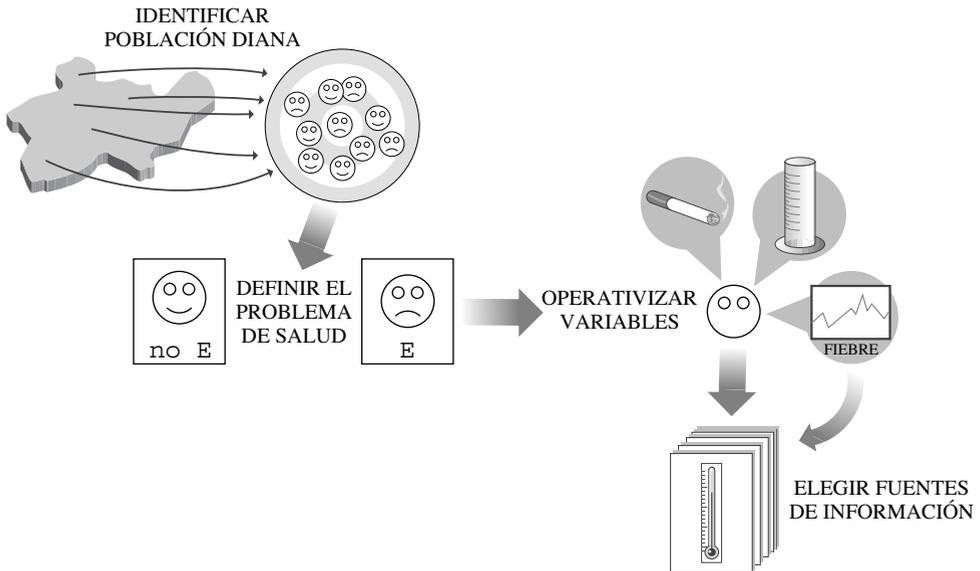


Fig. I.25- Etapas en el diseño de un estudio.

En primer lugar es necesario identificar la *población diana*. No necesariamente ha de ser población general. Podría existir interés en investigar la situación de cierto tipo de enfermos y la distribución en ellos de algunas características. O bien, el comportamiento de algunos servicios hospitalarios en relación a la calidad de la prestación que ofrecen. Todas estas situaciones llevan a la caracterización de distintas poblaciones diana.

De la misma forma, es imprescindible en cualquier diseño, la definición del *problema de salud* objeto de estudio. Sería demasiado vaga, por ejemplo, una descripción de casos de cáncer génito-urinario. Claramente se engloba con esta definición un conjunto de neoplasias demasiado diferentes entre sí, y se necesitaría una mayor especificación de los procesos que se pretende describir.

La fase de *operativización de variables* resulta en ocasiones especialmente vulnerable a 'lo obvio'. Un caso frecuente es el del nivel socioeconómico. En muchos estudios es difícil entender si los autores miden esa variable según ingresos familiares, nivel de estudios alcanzados, profesión del cabeza de familia, o simple impresión personal. Situaciones similares se encuentran en relación al estadiaje de tumores, nivel de gravedad de los pacientes e incluso en el manejo de parámetros bioquímicos. Es preciso entender además, que el carácter de las variables (continuas, categóricas, etc) determinará la indicación del tipo de análisis estadístico, y éso afecta tanto a las variables independientes (tipo de exposición) como a las dependientes (efecto o problema en estudio). Estos temas se tratan en el capítulo II.D.

La elección de las *fuentes de información* tiene importancia, por cuanto la disponibilidad de los datos puede ser determinante en relación al diseño que se pueda llevar a cabo. Así, si los datos están ya disponibles y son de buena calidad, sería viable el plantearse un estudio de tipo ecológico, usando datos secundarios. Si esa clase de información no existe, habrá que buscar datos primarios, usando cuestionarios o revisión de historias clínicas, por ejemplo.

Un punto de gran ayuda en la fase de diseño del estudio, es la decisión sobre los *indicadores y tablas* que se pretenden obtener en el análisis. Su utilidad práctica inmediata es la selección de las variables que deben incluirse en el protocolo, identificando cuáles sirven a los objetivos planteados y cuáles son innecesarias.

2. Informes de casos y series de casos

Esta aproximación descriptiva se basa en la experiencia clínica de un paciente o un grupo de pacientes con un diagnóstico similar. Es posible que se identifique una característica inusual de la enfermedad o de los antecedentes del o los enfermos. Existen ejemplos en la literatura en los que informes de este tipo lanzaron una pri-

mera hipótesis que después se verificó, o fueron la primera señal de alarma respecto a la aparición de nuevos síndromes o epidemias.

En 1961, se publicó un informe de un caso en el que una mujer premenopáusic¹⁹ desarrolló tromboembolismo pulmonar semanas después de comenzar el uso de anticonceptivos orales para el tratamiento de una endometriosis. Alrededor de 1981, en la ciudad de Los Angeles, se observó un número extrañamente alto de ingresos por Sarcoma de Kaposi y, en otros casos, de neumonía por *Neumocystis Carinii*, en personas jóvenes previamente sanas. Había un denominador común en ese exceso de diagnósticos: hombres de hábitos homosexuales, de ámbito urbano, y sin experiencia previa de inmunosupresión²⁰.

En 1959, en la ciudad santa de Meknes, en Marruecos, se detectó un raro hecho: en un corto periodo de tiempo apareció un número alto de casos de una enfermedad paralizante de origen desconocido. El médico responsable de los dispensarios dió la pista a los consultores que la OMS había enviado: había visto muestras de aceite comestible «tan oscuras como un aceite usado de motor», y algunos enfermos creían que ese aceite era el responsable de la enfermedad. La investigación epidemiológica posterior reveló que se había comercializado un aceite *de oliva* para consumo humano con sólo un 33% de aceite vegetal. Se encontró que contenía un lubricante artificial con fosfatos de cresilo; ese componente se usaba para resistir las altas temperaturas a las que se exponen los aceites usados para lubricar los turborreactores²¹. Un caso parecido ocurrió en España a comienzos de 1981²² y llevó a la identificación posterior del Síndrome del Aceite Tóxico español.

Aún cuando podrían citarse muchos otros ejemplos, y se calcula que en España las publicaciones de este tipo suponen entre el 50 y el 80% de la producción científica médica²³, su utilidad tiene una serie de limitaciones. En principio no sirven para probar hipótesis estadísticas, dado que representan la experiencia de una sólo persona o un pequeño número de ellas²⁴. Desde un punto de vista epidemiológico presentan un problema crucial: la inexistencia de algún grupo que pueda servir como comparación o control.

3. Estudios ecológicos

Este tipo de diseños presenta una particularidad única en el conjunto de aproximaciones posibles a la descripción de los problemas de Salud. La unidad de análisis es, como se ha comentado, grupos de individuos y no individuos en sí. Requieren la existencia y disponibilidad de datos agregados, ambientales o globales, algo que hace pocos años resultaba difícil de conseguir. La progresiva informatización de datos secundarios (ya existentes) está facilitando una enorme accesibilidad, y por lo tanto, la proliferación de análisis de tipo ecológico. Una ventaja indirecta es la existencia de

restricciones relativas a la confidencialidad de las bases de datos, cada vez más demandada por las sociedades avanzadas. Esto impide en muchos casos el uso de información individualizada²⁵ y obliga al uso de datos agregados.

Los diseños de tipo ecológico usan medidas que relacionan características de poblaciones en su conjunto, y permiten describir los problemas de salud en relación a elementos relacionados con características de las personas, el lugar y el tiempo. Esto permite cumplir objetivos tales como la rápida generación de hipótesis etiológicas, así como una fácil aproximación a la evaluación de programas preventivos o de intervención.

Entre sus ventajas más relevantes se puede citar su rapidez y bajo coste, ya que utilizan información disponible, además de la posibilidad de comparar grupos y zonas que necesitarían una enorme inversión de recursos si se usara otro tipo de diseño. Se acepta que pueden ser un primer paso en la investigación de una posible relación entre exposiciones y efectos variados. Sin embargo, existe una cierta controversia²⁶ sobre la validez de su uso en la verificación de hipótesis, sobre todo cuando se trata de evaluación de intervenciones.

Una clasificación de los diseños ecológicos los presenta en tres grandes grupos:

1.-Descriptivos o de mapas: comparan incidencia o prevalencia de un determinado problema en varias áreas, buscando la detección de patrones de tipo **geográfico**. Un buen ejemplo lo constituyen los denominados ‘Atlas de Cáncer’ (Fig. I.26). Pueden manejarse datos de incidencia o de mortalidad y en principio, se contrastan los resultados según áreas de tipo administrativo, como países, provincias o zonas hospitalarias de influencia. La tendencia de la mortalidad por cáncer de próstata a seguir un patrón de distribución costa-interior en Andalucía, por ejemplo, da pie a la generación de hipótesis sobre factores genéticos, dietéticos o culturales y de hábitos de vida²⁷.

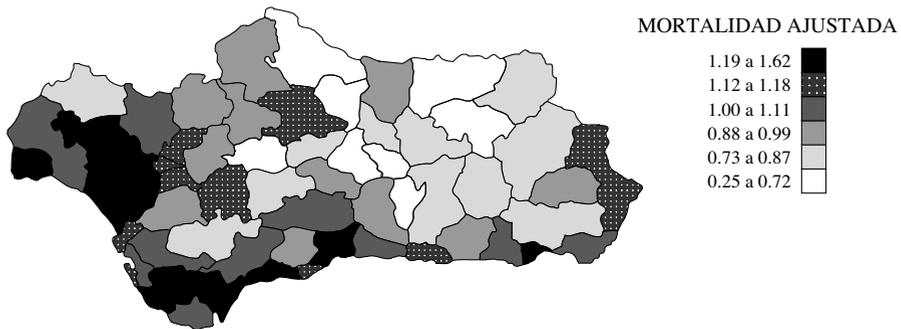


Fig. I.26- Mortalidad por tumor maligno de la próstata. Andalucía 1980-1984. (Distritos sanitarios)²⁷.

2.-Estudios de **Series Temporales**: describen el comportamiento de los problemas a lo largo del tiempo, ya sea buscando patrones estacionales, así como tendencias o cambios de tendencias a más largo plazo. Por ejemplo, el análisis de la mortalidad por cáncer de próstata en España muestra un incremento del 2% en promedio desde la década de los 50 a final de los 70²⁸. El aumento global durante todo ese periodo fue del 64%, una tendencia ascendente significativa. El aumento se produjo en dos fases: en la primera, de 1955 a 1970, la mortalidad se incrementó un 3.3% por año. En la segunda fase, hasta los 80, el incremento se ralentizó (0.07% por año). La tendencia ascendente se mantuvo hasta 1990. Su interés estriba en la posibilidad de predecir comportamientos y necesidades de atención sanitaria en el futuro (Fig. I.27).

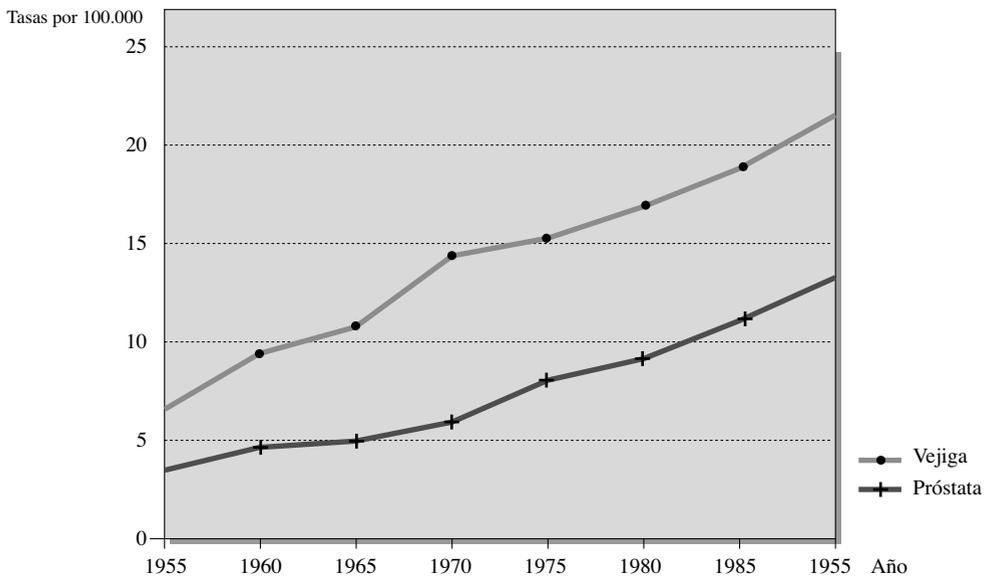


Fig. I.27- Tendencias de mortalidad por cáncer de vejiga y próstata. España, 1955-1990.

Fuente: Movimiento Natural de la Población.

Elaboración: Registro de Cáncer de Granada.

3.-Comparación entre grupos: estudios de **Correlación Ecológica**. Analizan el grado de relación entre niveles promedio de exposición y el comportamiento del efecto hipotético. Los grupos a comparar vienen definidos precisamente por la información disponible sobre los niveles de exposición. Por ejemplo, es posible describir la asociación del consumo promedio de tabaco y la mortalidad por infarto agudo de miocardio²⁹ en distintas zonas o en diferentes momentos en el tiempo. En España, el coeficiente de correlación entre ambos es de 0.94, enormemente alto (Fig. I.28). Se ha utilizado incluso esta aproximación para evaluar el impacto de intervenciones preventivas tales como programas de detección precoz de enfermedades ('screening')²⁶.

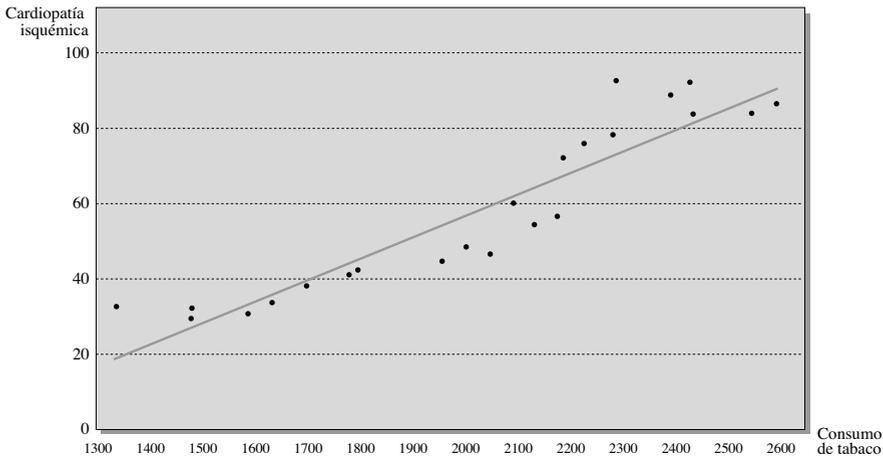


Fig. 1.28-Relación entre consumo per cápita de tabaco y mortalidad por cardiopatía isquémica en hombres. Tasas ajustadas por 100.000 habitantes. (España 1951-1979)²⁹.

Sea cual sea el tipo de diseño ecológico, sus propios puntos fuertes marcan sus debilidades. En el caso de la evaluación de programas de detección precoz de alguna enfermedad, por ejemplo, puede encontrarse una asociación positiva, y estadísticamente significativa, entre la cobertura del programa y el declive, años después, de la mortalidad específica por esa causa³⁰. Sin embargo, al usar datos agregados, no se demuestra definitivamente que la evitación de muertes se haya producido en el grupo sometido al 'screening'. Sólo se manejan datos promedio.

Por otro lado, el mismo hecho de manejar datos promedio puede estar ocultando algún factor 'distorsionante' o confundente. Es conocida la asociación entre consumo de carne y la mayor frecuencia de problemas como el infarto, el cáncer de colon o el de vejiga urinaria. Pero el consumo de carne podría estar asociándose a alguna otra variable, como un descenso en el aporte dietético de fibras, hábitos de vida más sedentarios y otros. Incluso, podría encontrarse relación con el número de horas que se pierden viendo la televisión. Y es posible que esos otros datos no estén disponibles para el estudio en cuestión.

Este tipo de situaciones tiene plenamente abierta la puerta en estudios ecológicos, y a veces la solución es bien difícil. Se ha llegado a describir una clara relación, lineal además, entre el porcentaje de población protestante y la tasa de suicidios en la Prusia de finales del siglo XIX, en un estudio citado por Morgenstern⁷ (Fig. 1.29). Podría ser cierto que las creencias de los protestantes aumentarían su riesgo de suicidio. Sin embargo, podría ocurrir que los suicidios fueran más frecuentes en la población católica conforme menor sea su número y se encuentre en una situación más marginal.

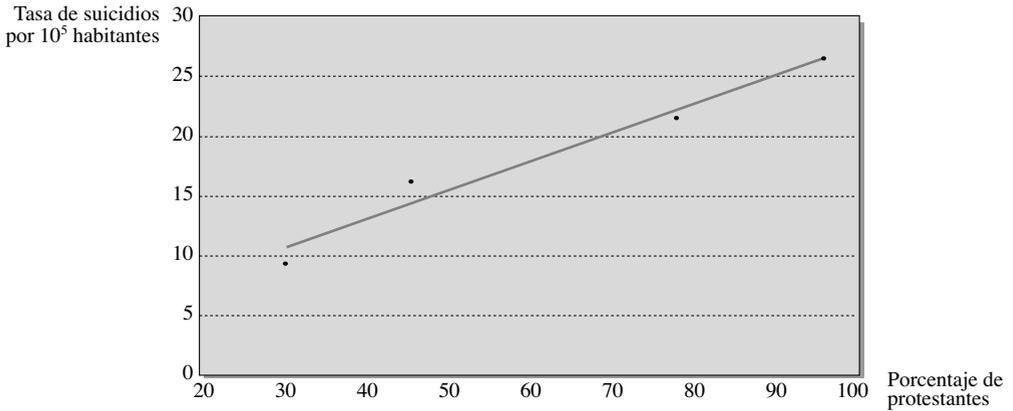


Fig. I.29- Tasa de suicidios (por 10⁵ habitantes) según composición religiosa en 4 provincias de Prusia. 1883-1890. Fuente: Morgenstern⁷.

Más aún, es posible demostrar que la disminución en el número de nidos de cigüeña es capaz de predecir un descenso en la natalidad, cuando no se tiene en cuenta el grado de urbanización o de industrialización. Se conoce como **Falacia Ecológica** a estos errores, derivados fundamentalmente del intento de extrapolar hallazgos de tipo ecológico a situaciones individuales, o cuando resultados provenientes de ciertos subgrupos se infieren a la totalidad de la población que los origina.

4. Estudios transversales (prevalencia)

En los estudios de este tipo, el investigador pretende evaluar simultáneamente, en una determinada población diana, el estado de exposición y la situación o el comportamiento de un problema de Salud dado (Fig. I.30). Suponen un paso más en relación a los diseños previamente comentados. No se trata de describir simplemente qué ocurre en una serie de casos, y además se dispone de información sobre todos los individuos investigados, o una muestra de ellos.

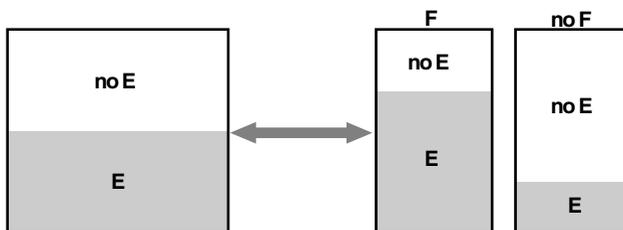


Fig. I.30- Estudio transversal. Prevalencia de un efecto (E) en expuestos (F) y no expuestos (no F).

Es decir, se supone una población diana bien definida y se trata de describir qué ocurre en su interior. En general se intenta conocer la situación puntual, en un momento dado en el tiempo, aunque en ocasiones la recogida de la información requiere un tiempo que puede ir de días a meses. Sin embargo, conceptualmente, se describe la situación en un punto determinado en el tiempo. Por ejemplo, qué proporción de individuos mayores de 50 años presentan algún tipo de hipertrofia prostática. O cuál es la prevalencia de enfermos con algún grado de insuficiencia renal crónica; aunque en este último caso se presupone la existencia de un sistema rutinario de registro.

Proporcionan datos útiles desde el punto de vista de la planificación y gestión de los sistemas de prestación de servicios sanitarios. El interés está no sólo en la identificación de necesidades de provisión de servicios. Se persigue además la detección de grupos más vulnerables o en mayor riesgo, que requerirán, en un sistema de prestación regido con criterios equitativos, de una atención *desigual*, es decir, preferencial.

Por otra parte, son diseños que permiten cruzar información acerca de la prevalencia de problemas de Salud con datos sobre la distribución de factores de exposición. Si además se puede comparar esa descripción con lo que ocurre en distintos grupos, las diferencias que se encuentren servirán para generar hipótesis, útiles en el diseño ulterior de estudios analíticos.

El desarrollo de estudios transversales requiere por lo general el uso de muestreo probabilístico o aleatorio. Los distintos tipos de muestreo aleatorio se tratan en otra parte de esta monografía (parte II). Desde el punto de vista del diseño, interesa recalcar la posibilidad de *comparar* la prevalencia de una situación (sea prevalencia de un problema relacionado con la salud o prevalencia de cierto factor de riesgo) en diferentes grupos y subgrupos, o entre diferentes poblaciones diana.

Como se compara simultáneamente exposición y presencia del efecto, surge una dificultad. A priori, debería pensarse que si esa exposición se comporta como factor de riesgo, las personas en las que se encuentre el problema de salud deberían presentar con mayor frecuencia la presencia de la exposición. Por ejemplo, cualquiera esperaría encontrar en pacientes con diabetes tipo II un mayor porcentaje de obesos.

Pero ¿qué ocurre si no es así? Un estudio de prevalencia podría describir a ese tipo de pacientes como individuos con un peso para la talla cercano al ideal, y seguidores de una dieta equilibrada y variada, en comparación con la población normal. Esos hábitos no son, obviamente, factores relacionados con un incremento en la probabilidad de padecer diabetes tipo II. Indudablemente la existencia del diagnóstico ha producido en los pacientes una modificación de su conducta. Hallazgos similares es

posible encontrarlos en estudios sobre situación nutricional, medida con indicadores de tipo antropométrico, y dieta. En países desarrollados es frecuente encontrar niños obesos que siguen una dieta estrictamente hipocalórica³¹. La presencia del problema ha modificado la situación de exposición.

Una amenaza constante en cualquier estudio transversal es qué viene primero (lo que se conoce como el dilema del *huevo y la gallina*). Podría argumentarse que una buena anamnesis sería capaz de identificar esos cambios y corregir la mala clasificación de los individuos en estudio. Es cierto, pero eso no elimina otra dificultad. Si el problema en estudio tiene una alta tasa de letalidad, o bien provoca cambios en la localización de los afectados, una corrección en la recogida de la información sólo sería útil para los que permanecen localizables o *sobreviven* más.

Es decir, pueden estar sobre-representadas las personas con un mejor pronóstico o supervivientes y con ellas, los factores asociados a un comportamiento más benigno de la enfermedad. Es frecuente por ejemplo, el consejo médico de un cambio de residencia para niños con asma. En la medida en que eso pudiera llevarse a cabo, sería fácil detectar una mayor prevalencia de asma infantil en zonas del litoral. De la misma forma, si la presencia de los primeros síntomas de obstrucción crónica al flujo aéreo determinara el traslado de mineros a otros puestos de trabajo, se encontraría una baja prevalencia de silicosis en quienes bajan a la mina, en comparación con los que trabajan en superficie.

Otro problema frecuente en los diseños transversales es el que se refiere a la *selección* de las personas que van a participar. Idealmente está definida una población diana, y en la mayoría de los casos se procederá, como se ha dicho, a un muestreo probabilístico. El establecimiento de criterios de inclusión y exclusión a partir de esa definición previa de la población diana debería evitar la participación de voluntarios. Este tipo de personas podría alterar la representatividad de la muestra, ya que puede tratarse de personas especialmente preocupadas por su salud; en ese caso, la prevalencia encontrada podría no representar lo que ocurre en la población de la que procede la muestra. O bien, podría ser que quienes se ofrecen como voluntarios son personas con especiales riesgos o atención sanitaria deficiente, y el error iría entonces en sentido contrario.

El proceso de localización de las personas que componen la muestra hace que, en ocasiones, se omitan las personas difíciles de identificar. Y en su lugar reemplazarlas por otras. Esto, que sólo es aceptable si se muestreó por error, tiene como consecuencia un *desplazamiento* de la muestra, que empieza a perder representatividad. El problema será entonces, a qué población se pueden extrapolar los resultados.

Más difícil es la solución de los *no respondentes*, las personas que por distintas razones se niegan a participar en el estudio, contestar cuestionarios, etc. Si son sistemáticamente diferentes en algo de los que participan, de nuevo la muestra podría no ser suficientemente representativa. Un recurso frecuentemente usado es determinar a priori un tamaño muestral mayor que el mínimo necesario, y reemplazar las negativas. Sin embargo, aunque de esa forma se soluciona el asunto del tamaño de la muestra, y por lo tanto su potencia para detectar diferencias, no ocurre lo mismo con el *desplazamiento* mencionado. Y si en lugar de muestrear se hubiera intentado investigar a toda la población, el problema habría sido el mismo, ya que los no respondentes en la muestra representan a aquellos que se negarían a participar del total de la población.

5. Análisis

La medida de frecuencia más inmediata en cualquier estudio transversal es lógicamente la prevalencia puntual, notada como $P(E)$. Se trata de dividir el número de casos que presentan el problema por la población existente en ese momento (Fig. I.31).

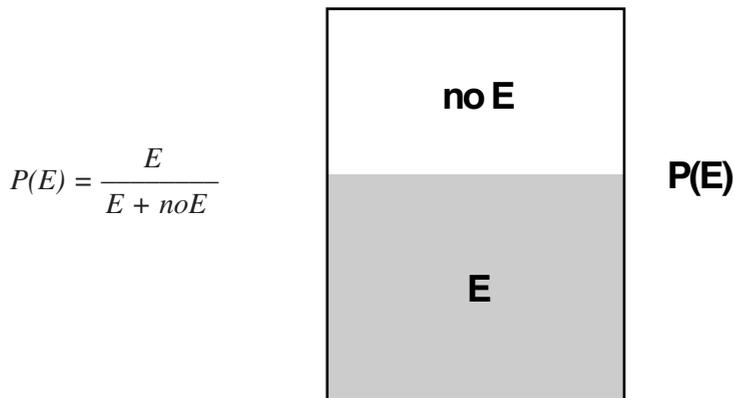


Fig. I.31- Prevalencia puntual.

Si además se pretende comparar entre distintos grupos caracterizados por la presencia o no de alguna exposición, se obtendrá entonces la prevalencia en los expuestos (F) y la de los no expuestos ($no F$). En este caso, interesa resumir las diferencias encontradas y puede calcularse la **Razón de Prevalencia**, como el cociente entre ambas. La tabla de la figura I.32 muestra los resultados al final de un estudio de prevalencia. En cada columna se sitúa el número de individuos que presentan el efecto, (E), o no (noE),

y en cada fila quiénes presentan exposición (*F*) y quiénes no (*no F*). La prevalencia en los expuestos se calcula dividiendo el número de personas enfermas en el grupo expuesto (*a*) por el total de expuestos (*a+b*). Idéntico cálculo puede hacerse para los no expuestos (*c/c+d*). La razón de prevalencia, finalmente, es el cociente entre ambas prevalencias (*RP*). Si se ha muestreado, ese cociente es un estimador de la verdadera razón de prevalencias que se da en la población.

	E	noE
F	<i>a</i>	<i>b</i>
noF	<i>c</i>	<i>d</i>

$$P(E)_F = \frac{a}{(a + b)} \quad P(E)_{noF} = \frac{c}{(c + d)}$$

$$RP = \frac{P(E)_F}{P(E)_{noF}} = \frac{\frac{a}{(a + b)}}{\frac{c}{(c + d)}}$$

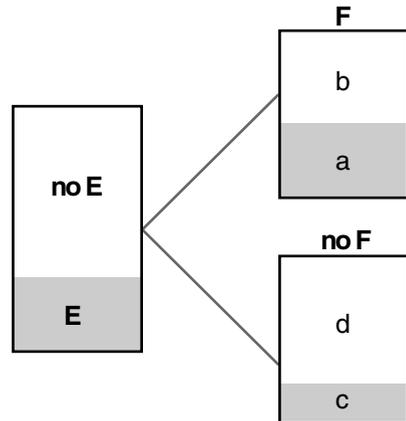


Fig. I.32- Razón de prevalencia (*RP*).

Otra aproximación consiste en considerar que se están comparando casos y controles (como controles se considerarían a los no casos), y calcular entonces las medidas de magnitud de asociación que se analizarán más adelante, como la **Odds Ratio (OR)** o **Razón de Oportunidades** (Capítulo I.F) de prevalencia.

Aunque, en principio, es más fácil entender un estudio transversal como uno en el que la variable dependiente es dicotómica (es decir, existe o no el problema en cada individuo), es posible que no siempre sea así. Por ejemplo, un estudio sobre la calidad de la prescripción farmacéutica en Atención Primaria³² encontró diferencias en los hábitos de prescripción de hepatoprotectores y otros medicamentos sin eficacia probada. Los datos procedían de una encuesta realizada a los médicos seleccionados en una muestra aleatoria. Se encontró una mayor prescripción en médicos que no asis-

tían habitualmente a sesiones clínicas y en los que dedicaban un menor tiempo de consulta por paciente. En este caso no puede hablarse estrictamente de *prevalencia* de mala prescripción, ya que la variable de efecto es continua (dosis diarias prescritas). O bien se trata la variable como continua, o bien se establece un punto de corte (basado en alguna argumentación) por encima del cual se considera que existe *mala prescripción*. En ese último caso puede calcularse la razón de prevalencia, así como la razón de oportunidades (*OR*).

6. Ventajas

-Los estudios transversales dan una descripción de la magnitud y el alcance del problema. Lo que conocemos como la *fotografía*, la situación en un momento dado.

-Proporcionan información basal para estudios de seguimiento posteriores, ya que permiten eliminar del seguimiento a los casos ya existentes, los *casos prevalentes* (Fig. I.33).



Fig. I.33- Corte transversal previo al seguimiento.

-Son útiles en evaluación de intervenciones sobre la población. Una combinación secuencial de estudios transversales permitiría evaluar el impacto sobre la prevalencia, si se espera una reducción en la incidencia o una mejoría en el pronóstico. Se ampliará este tema en el capítulo dedicado a diseños cuasiexperimentales (Fig. I.34).

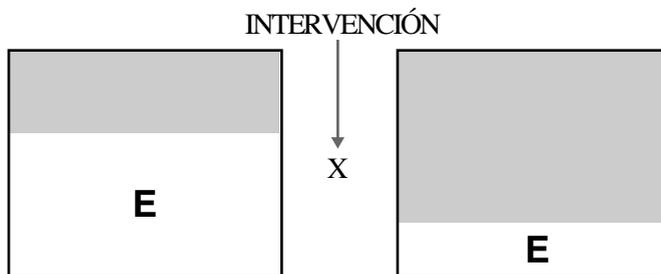


Fig. I.34- Combinación de dos cortes transversales para evaluar una intervención X.

-Pueden identificar casos (las personas con el efecto) y no casos (el resto), lo cual dá pie a un posible análisis como si se tratara de estudios de casos y controles (Fig. I.35). Sin embargo, esta opción no debería descuidar el hecho de manejar casos prevalentes, diagnosticados recientemente o más antiguos. Aún cuando pueda reconstruirse la experiencia previa de exposición, entre los casos detectados podrían estar aquellos que por la presencia de una determinada característica han disfrutado de una mejor supervivencia. Si es así, esa característica puede aparecer más frecuentemente en ellos, en comparación con los no casos. Entonces se concluiría erróneamente que la exposición ha jugado un papel de factor de riesgo, cuando en realidad pudo mejorar el pronóstico.

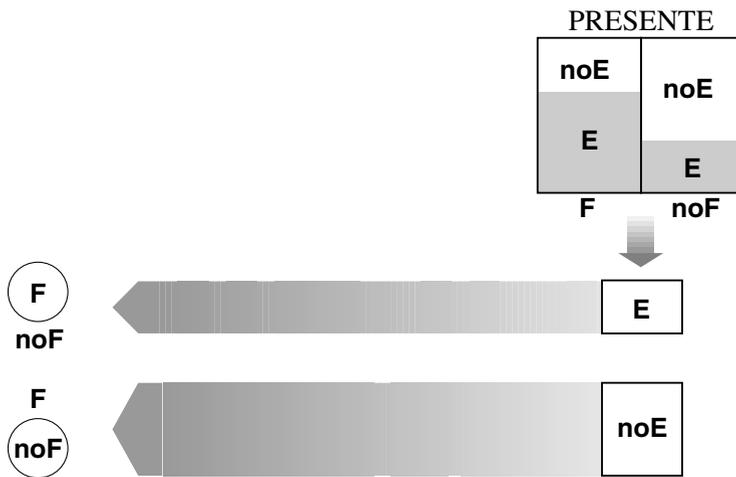


Fig. I.35- Análisis de un estudio transversal como estudio caso-control.

-Una ventaja no suficientemente valorada en la literatura es el bajo coste de este tipo de diseños en comparación con otros, especialmente los analíticos. Si existen pocos recursos son una opción a considerar frente a estudios más complejos, ya que se recolectan los datos de una sólo vez y no existe seguimiento. La única precaución que hay que tener es no olvidar sus limitaciones en cuanto a la posible verificación de hipótesis.

Aún cuando se han agrupado en la categoría de estudios meramente descriptivos, los estudios ecológicos y los transversales son en ocasiones la única opción, incluso cuando se intenta analizar determinados tipos de asociaciones. Los estudios transversales son muy útiles en las enfermedades crónicas no invalidantes, de larga duración, y de comienzo insidioso (como podría ser la hipertensión arterial o la hipertrofia benigna de próstata), o para el estudio de modificaciones de características fisiológicas (por ejemplo, cambios en los flujos expiratorios en la obstrucción crónica al flujo aéreo)³³.

Por otro lado, a pesar de la posibilidad siempre presente de la falacia ecológica, sólo los diseños ecológicos han sido capaces de detectar la magnitud de las pandemias modernas, como las relacionadas con la exposición a tabaco o a asbestos. Además, la evidencia ecológica siempre será un argumento a favor o en contra de hipótesis causales relacionadas con distintas exposiciones ³⁴.

E. ESTUDIOS DE SEGUIMIENTO

E. Perea-Milla López

1. Introducción

Son aproximaciones de tipo observacional, es decir, no experimentales. Sin embargo representan una posibilidad potente para la verificación de hipótesis, y muchas veces las evidencias obtenidas a partir de ellos permitirán tomar decisiones sin necesidad de acudir a la experimentación.

Se conocen como estudios de **cohorte** o, simplemente, estudios de **seguimiento**. Aunque en algunos textos se usa la denominación de *prospectivos*, esa terminología introduce cierta confusión y no es deseable, como veremos después. La idea central en este tipo de diseños es la observación a lo largo del tiempo de personas en las que el efecto que se investiga no está inicialmente presente. Se pretende comparar la aparición de nuevos casos (*E*) en el transcurso del tiempo, en grupos con diferentes niveles de exposición; en el caso más simple, un grupo expuesto (*F*) y otro (*no F*) en el que no existe exposición (fig. 36). Este último grupo es el que se usa como referencia. Como puede verse, la principal aportación que se introduce es el carácter longitudinal del diseño en el tiempo, algo inexistente en los estudios transversales.



Fig. 1.36- Estudio de seguimiento.

Cada grupo que se sigue se conoce como *cohorte*. El término proviene de la organización de las legiones romanas en compañías o cohortes en las que sus componentes tenían en común la procedencia de las mismas aldeas o comarcas. En los estudios de seguimiento, los individuos que componen cada cohorte tienen en común la existencia o no de un nivel dado de exposición. La población diana puede ser población sana, en la que es posible observar la aparición de daños a la salud como enfermedad, muerte u otros. También puede seguirse a poblaciones enfermas, en cuyo caso el efecto a investigar puede consistir en curaciones, recurrencias, mejorías o cualquier otro de interés tras el diagnóstico o desde el comienzo de los primeros síntomas. En cualquier caso, el propósito es reunir un número suficiente de personas *en riesgo* de que ocurra algo, y verificar si éso ocurre en un plazo más o menos lejano.

Ya se ha comentado la similitud de un estudio de seguimiento con otro experimental. La diferencia está en que aquí no existe asignación aleatoria de las personas a un grupo u otro. El investigador se limita a constatar si existe exposición, por las razones que sean, y no tiene control sobre la asignación ni interviene sobre la presencia de la exposición. Podría decirse que sólo *espera*. Si su hipótesis es cierta, la aparición de nuevos casos, la *incidencia*, será mayor en el grupo expuesto en comparación con el no expuesto.

No siempre un estudio de seguimiento requiere una actitud expectante durante un tiempo que podría durar años. Si el seguimiento de las cohortes se inicia en el momento en que comienza el estudio, entonces sí es inevitable. Aún más, habrá que ser cuidadoso en la medición del tiempo de exposición, que pudo estar presente antes de que el estudio se iniciara. En este caso suele hablarse de seguimientos *prospectivos*, ya que se vá hacia delante en el tiempo (Fig. I.37).

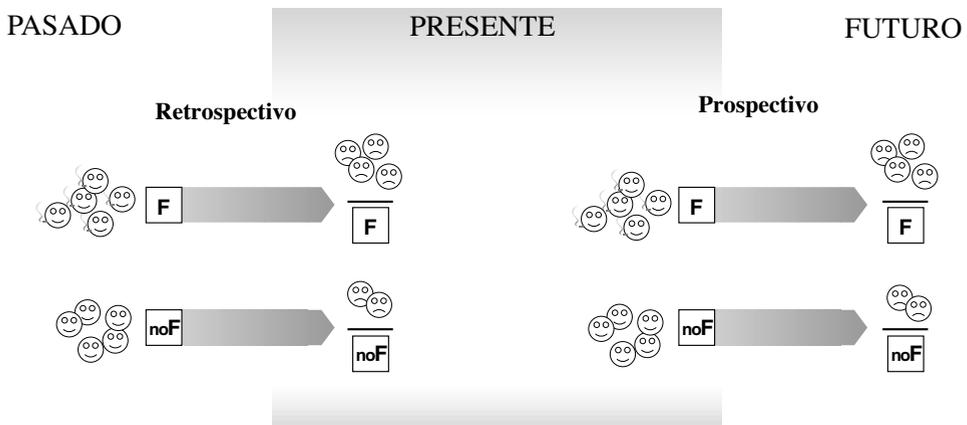


Fig. I.37- Estudios de seguimiento prospectivos y retrospectivos.

Existe otra posibilidad, y es aprovechar información pre-existente. Sería el caso en el que hubiera información válida previa, como historias clínicas bien cumplimentadas o buenos registros en servicios de medicina laboral, por ejemplo. Entonces, podría reconstruirse la experiencia previa de las cohortes, y usar esa información para conocer si, en grupos con niveles diferentes de exposición, hubo diferencias en la aparición del fenómeno en estudio.

Lógicamente la información histórica ha de ser suficientemente completa, válida y fiable. En el caso de registros de exposición laboral, la localización del puesto habitual de trabajo puede ser crucial para determinar si hubo o no exposición y en qué grado. Deberían conocerse los criterios diagnósticos utilizados, si hubo mayor vigilancia en unos trabajadores que en otros, etc. Cuando todo esto está disponible, es posible llevar a cabo lo que se conoce como estudio de seguimiento **histórico**, o *retrospectivo*.

Esta última denominación podría llevar a confusión, ya que se emplea también para los estudios caso-control (en los que el concepto '*retrospectivo*' indica que se parte del efecto y se investiga hacia atrás sobre la exposición). En los seguimientos de tipo histórico, ya que se parte de la exposición y se observa el efecto, una denominación adecuada sería estudios *prospectivos* (de la causa al efecto). Pero como además están localizados en el pasado, podría añadirse que son *retrospectivos*. Si además hubiera interés por continuarlos en adelante, el diseño final sería un estudio *prospectivo*, *retrospectivo* y *prospectivo*. Este galimatías es el que aconseja olvidar la palabra en cuestión. En adelante la usaremos sólo en un sentido estrictamente cronológico, según la situación del estudio en el tiempo.

2. Definición de las cohortes

La población diana que se va a seguir se identifica como población en estudio ('*study population*'), se muestree o no. Es posible partir de una población en estudio **fija** o **cerrada** (Fig. I.38). Esto quiere decir que se reúne un conjunto de personas, expuestas o no, y se piensa seguirlas para observar la aparición del o de los fenómenos de interés. No se admiten nuevas incorporaciones, una vez terminado el periodo de formación de la cohorte. Esto permite estimar la probabilidad de aparición del efecto en estudio tras un tiempo más o menos largo, que debe explicitarse. Por ejemplo, riesgo de presentación de complicaciones por infecciones oportunistas tras un trasplante renal, en un grupo de pacientes de cierta edad. En este caso, como se parte de una cohorte fija, es posible calcular la incidencia como **incidencia acumulada**. Simplemente se divide el número de nuevos casos por la población libre del efecto al comienzo del seguimiento, y esto puede hacerse en la cohorte expuesta y en la no expuesta.

Una cohorte cerrada no se sigue necesariamente *en tiempo real*. Por ejemplo, un estudio de seguimiento de pacientes diagnosticados de melanoma cutáneo, recogidos en el Registro de Cáncer de Granada³⁵. Los casos se diagnosticaron entre los años 1985 y 1989. Como fecha de inicio de las observaciones (el *tiempo cero* para el seguimiento) se consideró la fecha del primer diagnóstico. El seguimiento de la información se realizó de forma activa utilizando como fuente de datos las historias clínicas, contactos telefónicos y por correo con las familias, y los registros civiles de los ayuntamientos. Se pretendía comparar la mortalidad (efecto) según sexo y edad en los 3 primeros años desde el momento del diagnóstico. Ajustando por tipo histológico y nivel de Clark (grado de invasión), se encontró que la influencia del sexo sobre el pronóstico dependía de la edad de los pacientes.

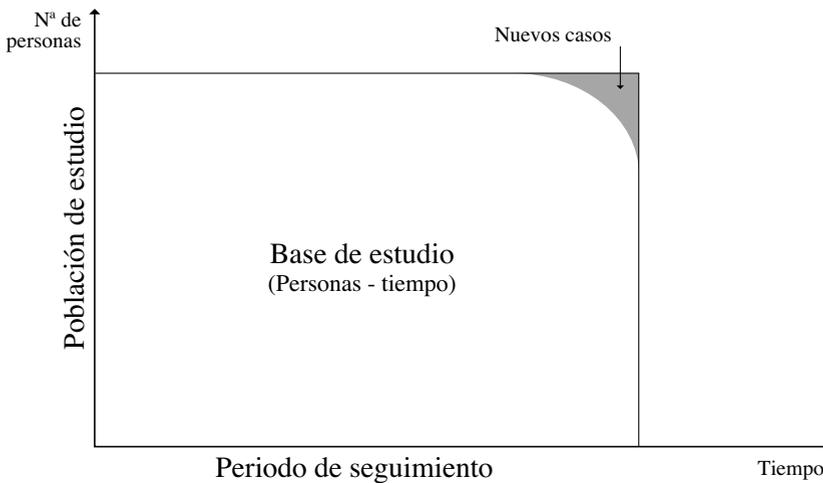


Fig. I.38- Población en estudio y base del estudio. Fuente: Norell S⁴⁴.

Por otro lado, en una cohorte cerrada es posible conocer el tiempo durante el que cada participante permanece libre del efecto. Entonces puede estimarse la cantidad de *tiempo en riesgo* observado en expuestos y no expuestos. Esa combinación de personas y tiempo se denomina **base del estudio** (*'study base'*) (Fig. I.38). El cociente en cada cohorte del número de nuevos casos aparecidos dividido por el sumatorio de *personas-tiempo*, o *base del estudio*, permite medir la incidencia como **densidad de incidencia**, o **tasa de incidencia** (Fig. I.39). Como se verá más adelante, la comparación de la incidencia de ambas cohortes es el fundamento del análisis en estudios de este tipo. Y esa comparación puede hacerse tanto si se dispone de incidencia acumulada como de densidad de incidencia.

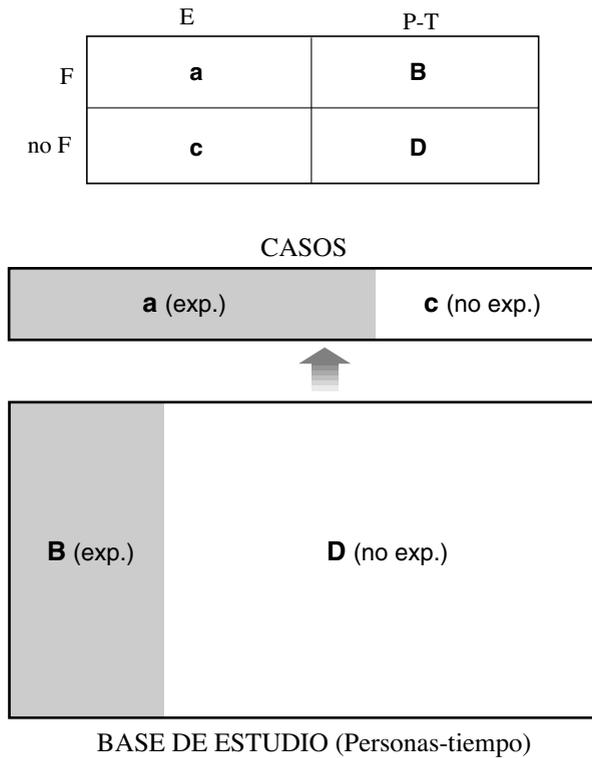


Fig. I.39- Los casos se originan en una población en estudio o en una base de estudio (suma de Personas-Tiempo o P-T) durante el periodo de estudio. Al final del seguimiento pueden resumirse los resultados en una tabla 2x2.
Fuente: Norell S⁴⁴.

Sin embargo, no siempre es ésa la situación. Puede ocurrir que las cohortes en seguimiento experimenten entradas y salidas. Puede imaginarse el seguimiento de una ciudad entera. Conforme avance el tiempo habrá gente que emigre, muera, nuevos niños que van naciendo y gente que viene a vivir. En ese caso no se dispone de una población fija, inicialmente libre del problema, y no es posible calcular la incidencia acumulada. Es el caso de una **cohorte abierta**. Aquí sólo queda la posibilidad de medir la **base de estudio**, y calcular densidad de incidencia. Si las entradas compensan a las salidas, entonces es una población *dinámica estable*. Por ejemplo, un estudio sobre la infección hospitalaria podría considerar los nuevos casos que aparecen durante un año, comparando servicios quirúrgicos y no quirúrgicos. Desde el momento del ingreso, cada paciente está en riesgo de adquirir una infección de ese tipo. Podría seguirse la aparición de nuevos casos *en tiempo real*. La suma de tiempos en riesgo individuales proporciona el denominador para el cálculo de la base del estudio (Capítulo I.C).

3. Identificación de las cohortes

La elección de los grupos a comparar depende de los investigadores, aunque no siempre³⁶. Una primera posibilidad es estudiar cohortes porque son más *fáciles de seguir*. Sería el caso de los médicos británicos seguidos por Doll y Hill para evaluar la influencia del tabaco sobre el riesgo de padecer cáncer de pulmón³⁷. En EEUU es frecuente el uso de personas que han suscrito un seguro de asistencia sanitaria. En nuestro ámbito una buena opción para conseguir cohortes fáciles de seguir es, por ejemplo, la de los donantes de sangre³⁸. En este caso se trata de subdividir la cohorte original en distintos subgrupos según el nivel de exposición. El grupo de comparación es **interno**, se obtiene de la propia cohorte.

Otra posibilidad es seguir a determinadas cohortes porque *han sido expuestas* a un agente hipotético. Por ejemplo, los supervivientes de las bombas atómicas de Hiroshima y Nagasaki. O los afectados por accidentes nucleares como el de Chernobil, en los ochenta, o el de la Isla de las Tres Millas, en la década de los setenta. Más cercano a nosotros, pudo haberse iniciado un seguimiento de los hipotéticos expuestos en 1966 cuando un accidente hizo caer algunas cabezas nucleares portadoras de plutonio en Palomares (Almería). En general, la aparición de una exposición obliga a realizar el seguimiento. El problema en estas situaciones es que se dispone de una cohorte expuesta, pero se necesita un grupo de comparación **externo**.

Una solución es usar la población general como *cohorte no expuesta*. Tiene la ventaja de aportar la incidencia que se esperaría en la cohorte expuesta si no hubiera existido la exposición. Sin embargo, puede haber problemas de comparabilidad; además, si la exposición es frecuente, su caracterización de 'no exposición' pierde fuerza. Y la vigilancia a la que están sometidas las cohortes puede ser diferente. Otra opción es utilizar una cohorte distinta como cohorte no expuesta, lo más similar posible a la expuesta. Por ejemplo, trabajadores de otras fábricas, en el caso de riesgos laborales. En ocasiones, la dificultad para encontrar una cohorte de comparación adecuada obliga a comparar los distintos niveles de exposición con el grupo menos expuesto, que se convierte así en el grupo de referencia.

4. Formación de las cohortes

El primer paso es identificar a las personas *susceptibles* de padecer la enfermedad o sufrir el efecto que se estudia (*E*). En algunas situaciones es relativamente fácil. En enfermedades transmisibles, por ejemplo, puede bastar la detección de anticuerpos para detectar a los ya inmunizados. En otras ocasiones es necesario un estudio

de prevalencia previo para detectar los *casos prevalentes* (Fig. I.40). Si ya padecen el efecto, no están en riesgo de desarrollarlo, y por lo tanto hay que excluirlos del seguimiento.



Fig. I.40- Estudio de prevalencia previo al seguimiento.

5. Determinación de la exposición

Las técnicas o instrumentos de medición requieren unos niveles de validez y fiabilidad conocidos (Capítulo II.B). Pueden ser cuestionarios, reconocimientos periódicos, muestras biológicas o información recogida con anterioridad. Antes de iniciar el seguimiento es recomendable en general un periodo de pilotaje para adaptar los cuestionarios, homogeneizar la técnica y conseguir un buen entrenamiento de los observadores. El objetivo es obtener una clasificación adecuada de la exposición. Un problema añadido, según el tipo de cohorte de que se trate, es el cambio en los niveles de exposición a lo largo del tiempo.

6. Sesgos

Como no existe asignación aleatoria, cualquier estudio de seguimiento es más vulnerable a errores sistemáticos que un estudio aleatorio controlado. Este problema afectará en mayor medida al control de las posibles variables confundentes, que se tratan en otro lugar.

6.1 Sesgo de selección

Se produce cuando la cohorte que finalmente se consigue observar difiere de la cohorte original en algún aspecto relevante. Un aspecto que merece especial atención es el de la *autoselección*. Puede haberse decidido utilizar a los donantes de sangre,

aprovechando las campañas periódicas de recogida de sangre. Es posible que personas especialmente expuestas, o con una peor accesibilidad a los servicios sanitarios, intenten participar en los reconocimientos. Si sus características están de alguna forma relacionadas con la probabilidad de aparición del efecto, entonces su participación en el estudio modificaría las estimaciones de los riesgos. Por ejemplo, el Baltimore Longitudinal Study of Aging (BLSA)³⁹ utilizó voluntarios para conocer, entre otros efectos, la incidencia de hipertrofia benigna de próstata (HBP). Si la razón por la que los voluntarios se interesan por participar en el estudio se relaciona con la probabilidad de desarrollar HBP y con los factores de riesgo analizados, entonces las estimaciones del estudio podrían resultar sesgadas. Si no es así, la pregunta que queda por responder es en qué medida esos voluntarios son representativos del resto de la población general.

Los estudios de seguimiento en poblaciones laborales encuentran frecuentemente hallazgos paradójicos. Se puede observar, por ejemplo, que entre los mineros activos la mortalidad es menor que la de la población general. Es el efecto del *trabajador sano*. La causa es que la gente más sana es la que permanece activa, mientras que en la población general están contenidos los ancianos y los incapacitados, personas con un mayor riesgo de morir.

La mayor amenaza relacionada con la selección la constituyen los *abandonos* en los estudios de seguimiento. Aparte de su efecto sobre la potencia estadística, si las personas de las que se pierde información son sistemáticamente distintas a quienes permanecen en el estudio, y las diferencias están de alguna forma relacionadas con la aparición del efecto, es razonable esperar una modificación en las estimaciones.

6.2. Sesgo de clasificación

Ya se ha hablado sobre la importancia de clasificar adecuadamente los diferentes niveles de **exposición**. Múltiples razones pueden alterar ésto. Por ejemplo, un estudio que explorara la influencia de las relaciones sexuales fuera de la pareja habitual encontraría serias dificultades. Por razones de rechazo o aceptación social, sería posible clasificar como monógamos a muchos individuos que mantuvieran relaciones con más de una pareja sexual, o todo lo contrario. Si esa exposición estuviera realmente relacionada con el efecto que se estudia, entonces la cohorte ‘no expuesta’ contendría erróneamente personas realmente expuestas, y en los ‘expuestos’ habría personas que en realidad no estuvieron expuestas. En ese caso el efecto de la mala clasificación iría en el sentido de producir una mayor incidencia, *artificial*, en la cohorte no expuesta, y lo contrario en los expuestos. En consecuencia, unas diferencias realmente existentes quedarían *diluidas* por la mala clasificación (Fig. I.41). Se habla entonces de *mala clasificación no diferencial*, una sospecha a considerar en todo estudio no concluyente. Si el estudio, al contrario, encontró diferencias, la posible existencia de este

sesgo preocupa menos, ya que en todo caso las diferencias son al menos tan grandes como las encontradas, o mayores.

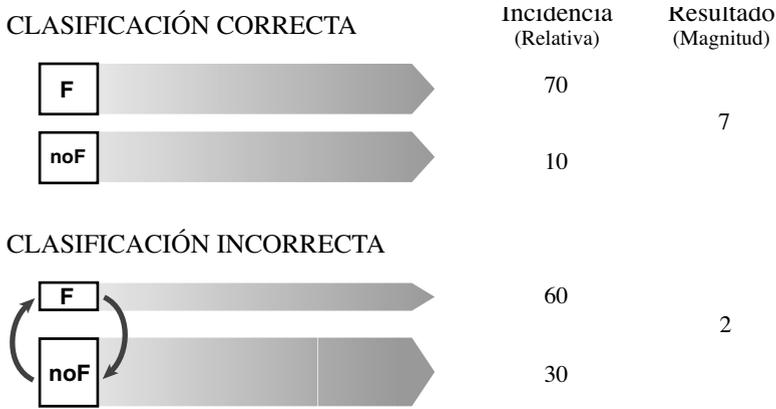


Fig. I.41- Clasificación incorrecta no diferencial.

La clasificación errónea de expuestos y no expuestos provoca que las diferencias sean menores cuando se compara en términos relativos la incidencia de ambas cohortes.

Quando el error en la clasificación se da en los expuestos y también en los no expuestos, entonces es no diferencial.¹⁴

Los errores en la clasificación pueden afectar también al **efecto**. Los pacientes sometidos a un nuevo procedimiento terapéutico están sujetos por lo general a una mayor vigilancia en su seguimiento. Si la vigilancia diagnóstica está relacionada con la existencia de la exposición, se puede esperar una mayor incidencia *detectada* en la cohorte expuesta. Puede imaginarse el trabajo de los servicios de Medicina Laboral en una central nuclear, en cuanto a la aparición de signos y síntomas de leucemia, por ejemplo. Si la cohorte de comparación es población general, podría darse el caso de enfermos no diagnosticados o deficientemente etiquetados. Entonces se encontraría una mayor incidencia en la cohorte expuesta, aún cuando la exposición no tuviera relación con ese efecto. En este caso la mala clasificación actúa en el sentido de *exagerar* diferencias, aún cuando no existieran. Se trata de una **mala clasificación diferencial**, porque el error ocurre sólo en los no expuestos (la población general) y no en ambos grupos. En cualquier estudio de seguimiento puede existir mala clasificación (diferencial o no diferencial) que afecte a la exposición o al efecto.

7. Análisis

La medida de frecuencia que se obtiene en un seguimiento es la incidencia, sea como incidencia acumulada (riesgo) o como densidad de incidencia (tasa). La comparación entre las cohortes implica, por tanto, una comparación entre incidencias

7. 1. Magnitud o fuerza de la asociación.

El interés aquí es conocer en qué medida la incidencia en los expuestos es mayor, menor o igual que en los no expuestos, que se usan como referencia.

a. Riesgo Relativo.

Una idea inmediata es dividir la incidencia en expuestos (I_F) por la incidencia en los no expuestos o grupo de referencia (I_{noF}); el resultado, la **razón de incidencias**, se conoce como **Riesgo Relativo (RR)**. Si se usa incidencia acumulada, el riesgo relativo es una razón de riesgos; en el caso de la densidad de incidencia se trata de una razón de tasas, aunque por costumbre se mantiene el mismo nombre.

Razón de riesgos.

(Tabla I.1). El riesgo relativo puede ser mayor que 1, cuando la incidencia es mayor en la cohorte expuesta. Si se trata de la aparición de una enfermedad, entonces la exposición está asociada con un aumento en la incidencia, en comparación con el grupo de referencia. Si es menor que la unidad, entonces la exposición se asocia con un efecto protector. Y si es igual a la unidad, no hay diferencias entre exposición o no exposición o, lo que es lo mismo, no hay evidencias de asociación.

Cuando se usa con incidencia acumulada, el riesgo relativo informa de cuántas veces se incrementa el riesgo de padecer el problema en los expuestos en relación a los no expuestos, en el caso de un factor de riesgo o de mal pronóstico. Es información útil para identificar incrementos de riesgo individuales si se produce la exposición.

	<i>E</i>	<i>no E</i>
<i>F</i>	<i>a</i>	<i>b</i>
<i>no F</i>	<i>c</i>	<i>d</i>

$$I_F = \frac{a}{(a+b)} \qquad I_{noF} = \frac{c}{(c+d)}$$

$$RR = \frac{I_F}{I_{noF}} = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}}$$

TABLA I.1- Razón de riesgos. Cociente entre incidencia en expuestos (I_F) e incidencia en no expuestos (I_{noF}).

Los datos que siguen a continuación (tabla I.2) provienen de un estudio en el que se siguió a 259 pacientes con tumor de Wilms que habían sido sometidos a radioterapia⁴⁰, para evaluar el riesgo de recidivas (E). Se les clasificó según el tipo de irradiación en función del tamaño del campo irradiado, clasificado como suficiente ($no F$) o insuficiente (F).

	E	$no E$
F	2	21
$no F$	2	234

$$I_F = \frac{2}{(2+21)} = 0.09 \qquad I_{noF} = \frac{2}{(2+234)} = 0.009$$

$$RR = \frac{I_F}{I_{noF}} = \frac{\frac{2}{(2+21)}}{\frac{2}{(2+234)}} = 10$$

TABLA I.2- Razón de Riesgos o Riesgo Relativo (riesgo de recidiva en casos de tumor de Wilms sometidos a radioterapia).

Los resultados del seguimiento indican que la incidencia acumulada de recidivas en los poco irradiados fue de 0.09. Es decir, el riesgo de sufrir una recidiva tras un periodo de tiempo X es del 9% en ese grupo. Sin embargo, ese mismo riesgo es menor en los radiados correctamente (9 por mil o 0.009), con lo cual el riesgo de recidiva se multiplica por 10 si se irradia insuficientemente a un paciente con tumor de Wilms. Aunque se trata en otro capítulo, puede señalarse que los resultados fueron estadísticamente significativos ($p=0.003$).

Razón de tasas.

(Tabla I.3). Cuando se usa con densidad de incidencia (nuevos casos, E , en relación a cantidad de personas-tiempo o $P-T$), el riesgo relativo estima el número de veces que aumenta la velocidad en el flujo de sanos a enfermos (fuerza de morbilidad en este caso) en el grupo expuesto (a/N_F) frente al de referencia, el no expuesto (c/N_{noF}).

	<i>E</i>	<i>P-T</i>
<i>F</i>	<i>a</i>	N_F
<i>no F</i>	<i>c</i>	N_{noF}

$$I_F = \frac{a}{N_F}$$

$$I_{noF} = \frac{c}{N_{noF}}$$

$$RR = RT = \frac{I_F}{I_{noF}} = \frac{\frac{a}{N_F}}{\frac{c}{N_{noF}}}$$

TABLA I.3- Razón de tasas. Cociente entre la tasa en expuestos y la tasa en no expuestos.

En el siguiente ejemplo hipotético, el seguimiento podría hacerse en un grupo de pacientes con hipertrofia benigna de próstata (HBP) desde el momento de su diagnóstico (sería el caso de una cohorte cerrada). El efecto en estudio pudo ser la indicación de cirugía, dependiendo de la situación inicial de los pacientes (según sintomatología, flujo urinario, tamaño de la próstata, etc).

Otra posibilidad sería seguir a pacientes sanos (sin HBP) de un colectivo determinado (por ejemplo, los donantes de sangre en un centro regional) y cuantificar la aparición de nuevos casos de hiperplasia mediante exámenes periódicos, para comparar la incidencia en distintos grupos de exposición; en este caso podría plantearse el seguimiento en tiempo real como una cohorte abierta, con personas que comienzan a donar sangre y otras que dejan de hacerlo. En el ejemplo (Tabla I.4), la tasa de incidencia en expuestos es de 0.0006, o lo que es lo mismo, 6.6 casos por cada 10.000 años, haciendo una simple regla de tres. Se interpreta como el número de nuevos casos aparecidos por unidad de tiempo, en este caso años. Por cada 10.000 años de exposición que se pudieran observar de personas expuestas al factor *F*, se producirían 6,6 nuevos casos de enfermedad.

	<i>E</i>	<i>P-T</i>
<i>F</i>	60	90.000
<i>no F</i>	12	96.000

$$I_F = \frac{60}{90.000} = 0.00066 \text{ año}^{-1} \quad I_{noF} = \frac{12}{96.000} = 0.00012 \text{ año}^{-1}$$

$$RR = RT = \frac{I_F}{I_{noF}} = \frac{\frac{60}{90.000}}{\frac{12}{96.000}} = 5.33$$

TABLA I.4- Cálculo de la Razón de Tasas.

En los no expuestos, la tasa de incidencia ha sido de 0.00012. De la misma forma, cada 10.000 años en riesgo producen 1.2 nuevos casos. Como puede verse, son incidencias realmente bajas, ya que cada año en riesgo no llega a producir ni siquiera un caso. La velocidad de aparición de nuevos casos es tan baja que el avance de la enfermedad se produce a razón de una fracción de persona por año.

La Razón de Tasas, cuyo valor es de 5.33, expresa que la velocidad de aparición de nuevos casos es más de 5 veces mayor en el grupo expuesto al factor F. El cual constituye un factor de riesgo, obviamente.

b. Odds Ratio o Razón de Oportunidades.

Otra aproximación posible a la estimación de la magnitud de la asociación es como sigue. Supongamos un factor que incrementa el riesgo de que aparezca una enfermedad. Si éso es así, se espera que al final del seguimiento la relación entre enfermos y no enfermos sea más *favorable* a los primeros (mayor proporción de enfermos) en la cohorte expuesta que en la no expuesta. Esa relación se puede expresar como un cociente entre proporción de enfermos y la de no enfermos en cada cohorte. Es un cociente entre una probabilidad (la de haber desarrollado la enfermedad) y su probabilidad complementaria (no haberla desarrollado), lo que en inglés se conoce como **odds** (*ventaja u oportunidad*, no hay una traducción clara). Se puede estimar entonces la *odds* de incidencia o de enfermedad en los expuestos, y la *odds* de incidencia

o enfermedad en los no expuestos (Tabla I.5). Como se ha dicho, es de esperar que la primera sea mayor que la del grupo de referencia. Y al tratarse de probabilidades, se entiende que se está haciendo una aproximación al caso de la incidencia acumulada y por tanto ésto es de aplicación en cohortes fijas o cerradas.

	<i>E</i>	<i>no E</i>
<i>F</i>	<i>a</i>	<i>b</i>
<i>no F</i>	<i>c</i>	<i>d</i>

$$Odds_i(F) = \frac{\frac{a}{(a+b)}}{\frac{b}{(a+b)}} = \frac{a}{b} \qquad Odds_i(no F) = \frac{\frac{c}{(c+d)}}{\frac{d}{(c+d)}} = \frac{c}{d}$$

$$O.R. = \frac{Odds_i(F)}{Odds_i(noF)} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a*d}{b*c}$$

TABLA I.5- Odds Ratio (OR) de incidencia. Cociente entre la odds de incidencia en expuestos, $Odds_i(F)$, y la odds de incidencia en no expuestos, $Odds_i(no F)$.

Por lo tanto, un cociente que relacionara ambas *odds* (**Odds ratio**, **OR** o **Razón de Oportunidades**), sería mayor que 1 para un factor de riesgo. Si fuera un factor protector, las cosas serían totalmente opuestas, y la *OR* sería inferior a la unidad. En caso de no haber asociación no hay razón para esperar diferencias en las dos cohortes, y la *OR* sería igual a 1. Un comportamiento idéntico al *RR*, como puede verse. Se puede verificar que la *OR* es igual a $(a*d)/(b*c)$, por lo que también se conoce como **Razón de Productos Cruzados**. Cuando la enfermedad es poco frecuente, y ocurre en la mayoría de las ocasiones, la *OR* dá valores similares a la Razón de Riesgos o Riesgo Relativo (*RR*).

7. 2. Medidas de impacto potencial.

Además de la magnitud de la asociación, interesa conocer qué cantidad del daño a la salud es atribuible a la existencia de la exposición. Si pudiera cuantificarse ésto, sería posible entonces estimar cuánto se podría reducir el daño si se pudiera eliminar la exposición, totalmente o en parte.

Riesgo Atribuible.

Con el mismo ejemplo de un factor asociado a un mayor riesgo de padecer la enfermedad, la situación al final del seguimiento sería la de una mayor proporción de nuevos casos en la cohorte expuesta (Fig. I.42). De forma visual, usando incidencia acumulada, aproximadamente dos tercios de los expuestos en la figura han enfermado ($I_F=0.66$), frente a sólo un tercio en el grupo de referencia ($I_{noF}=0.33$). El riesgo relativo es, pues, aproximadamente igual a 2 ($0.66/0.33$).

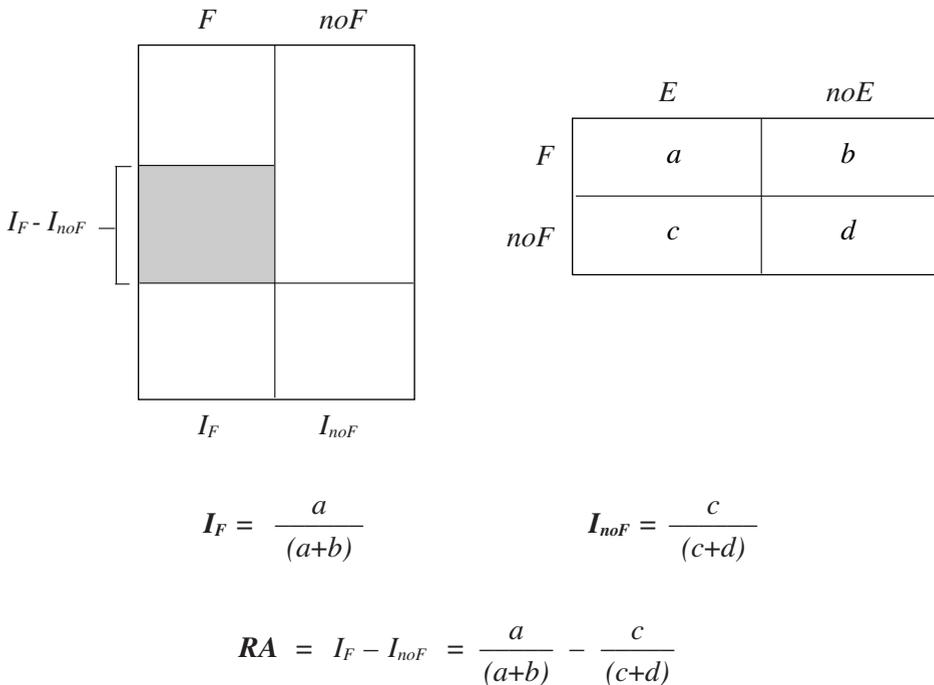


Fig. I.42- El impacto potencial: cohorte expuesta y no expuesta.

En esa situación una forma directa de cuantificar el exceso de incidencia observado en los expuestos es restar la incidencia en los no expuestos (I_{noF}) de la de los expuestos (I_F). Esa diferencia, conocida en general como **Riesgo Atribuible (RA)**, cuantifica el *exceso de riesgo* producido por la exposición, *en el grupo expuesto*, si se usa incidencia acumulada. Como es una resta entre dos incidencias, el resultado es otra incidencia. En el ejemplo de la figura, $0.66-0.33=0.33$. Es decir, en un tercio de los expuestos se ha desarrollado la enfermedad en exceso, y ello es atribuible a la presencia de la exposición en ese grupo. Dicho de otra forma, ésta es la cantidad de expuestos que no desarrollarían la enfermedad, la cantidad de enfermedad que podría evitarse, si desapareciera la exposición.

Con densidad de incidencia, se cuantifica el exceso en la tasa producido por la exposición; es decir, la *aceleración* en el proceso de enfermar introducida por la presencia de la exposición. La incidencia de enfermedad observada en la cohorte no expuesta es el nivel basal, la velocidad de aparición de nuevos casos que se esperaba encontrar si no existiera exposición.

Siguiendo con el mismo ejemplo, podría expresarse el exceso de riesgo (en el caso de la incidencia acumulada) como la proporción que representa respecto a la frecuencia de casos aparecidos en la cohorte expuesta. Es decir, ¿qué proporción de casos aparecidos en la cohorte expuesta se debe precisamente a la presencia del factor de riesgo en estudio?. Dicho de otra manera, ¿qué proporción del efecto se podría eliminar si desapareciera la exposición?. Se trata de construir un porcentaje que relacione el exceso de riesgo encontrado ($I_F - I_{noF}$) con el riesgo observado en los expuestos (I_F). Se está determinando la amplitud de la zona sombreada en la figura respecto a la incidencia *en la cohorte expuesta* (visualmente, es de aproximadamente el 50%). Esta cantidad recibe el nombre de **Proporción atribuible en expuestos (PA)**, y se expresa según la fórmula:

$$PA = \frac{I_F - I_{noF}}{I_F}$$

Con los datos de la gráfica anterior, la Proporción Atribuible en expuestos es, efectivamente del 50% (0.50). Es decir, se podría haber evitado la mitad de los casos que han aparecido en la cohorte expuesta si se hubiese evitado esa exposición:

$$PA = \frac{0.66 - 0.33}{0.66} = 0.50$$

Puede demostrarse que la expresión anterior es igual a la siguiente:

$$\begin{aligned} PA &= \frac{I_F - I_{noF}}{I_F} = \frac{I_F}{I_F} - \frac{I_{noF}}{I_F} = \\ &= 1 - \frac{I_{noF}}{I_F} = 1 - \frac{1}{RR} = \\ &= \frac{RR}{RR} - \frac{1}{RR} = \frac{RR - 1}{RR} \end{aligned}$$

En el ejemplo, asumiendo un Riesgo Relativo de 2 ($0.66/0.33=2$), la Proporción Atribuible en expuestos queda igual a un 50% [$(2-1)/2 = 0.50$], tal como ya se había visto. El interés de todas estas operaciones está en la posibilidad de calcular esta medida de impacto potencial en estudios de casos y controles, como se verá más adelante.

De la misma forma que para los expuestos, es posible calcular qué cantidad de daño a la salud sería evitable si se eliminara la exposición, pero teniendo en cuenta el total de casos aparecidos en una población, y no sólo en los expuestos. Esta otra medida de impacto potencial se conoce como **Proporción Atribuible Poblacional (PAP)**, aunque según los autores se encuentran denominaciones distintas, como *riesgo atribuible poblacional*, o *fracción etiológica*. Se puede demostrar que esta fracción es igual a la siguiente expresión:

$$PAP = \frac{F * (RR - 1)}{1 + (F * (RR - 1))}$$

donde F es la prevalencia del factor de exposición en la población (considerada ésta como la suma de expuestos y no expuestos). En el ejemplo anterior, en una población con la mitad de la gente expuesta (el tamaño de ambos grupos es igual) puede calcularse de forma visual que el exceso de riesgo observado en la cohorte expuesta supone aproximadamente un tercio de la incidencia total. Luego el valor que se obtenga de la PAP será de 0.33 (un 30%). De nuevo, el interés de esta formulación se hará extensible a los estudios de caso-control.

Como la eliminación total de una exposición puede no ser factible, existen procedimientos de cálculo para el caso de reducciones parciales, así como para el caso de análisis multivariante⁴¹.

No debe olvidarse que las medidas de magnitud y de impacto potencial calculadas son *estimadores* de la verdadera relación existente entre exposiciones y efectos. Por otra parte, se ha presentado un análisis simple, lo que se conoce como análisis *crudo*. No se ha tenido en cuenta la posible distorsión en los resultados que podría provocar una distribución desigual de alguna otra variable que actuara como *confundente*. La respuesta a estas cuestiones se trata en otras secciones.

F. ESTUDIOS DE CASOS Y CONTROLES

E. Perea-Milla López

1. Introducción

Suponen una aproximación más eficiente, en el terreno no experimental, a la verificación de hipótesis. Básicamente cumplen los mismos objetivos que los estudios de seguimiento, con un menor consumo de recursos y tiempo. Incluso son la alternativa cuando se trata de problemas poco frecuentes, o con períodos de inducción largos. Sin embargo son más vulnerables a la presencia de determinados errores sistemáticos o sesgos.

Clásicamente, la idea central es comparar personas en las que el efecto (enfermedad, muerte u otro) ha aparecido, los **casos**, con personas semejantes, los **controles**, en las que éso no se ha producido. El criterio de asignación es ahora la presencia o no del efecto. Suponiendo un factor de riesgo, es verosímil que los casos presenten con mayor frecuencia antecedentes de exposición que los controles (Fig. I.43). Podría llegarse entonces a conclusiones semejantes a las que proporcionaría un estudio de seguimiento. Sin embargo, como veremos más adelante, en realidad un estudio caso-control tiene enormes semejanzas con uno de cohortes, sólo que aquí se obtiene información sobre la distribución de la exposición a partir de una muestra de la *población en estudio* o de la *base de estudio*.

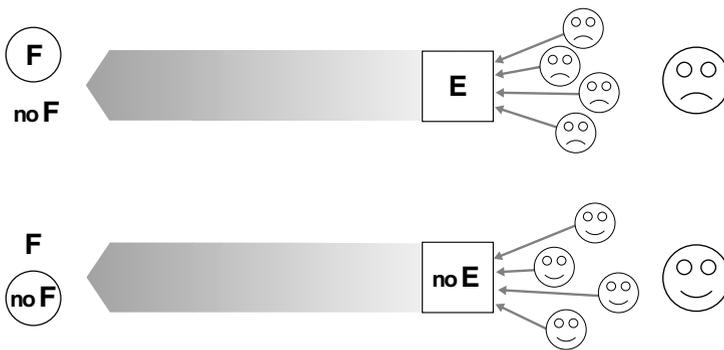


Fig. I.43- Estudios de casos y controles: una aproximación inicial.

Como ejemplo, se diseñó un estudio caso-control multicéntrico ⁴² para valorar el riesgo asociado a la ocupación, el consumo activo y pasivo de tabaco, la dieta, café, analgésicos, edulcorantes artificiales y los antecedentes de infección y litiasis urinaria.

Se incluyeron como casos todos los enfermos diagnosticados de carcinoma de vejiga urinaria o tumor de vejiga (Código 188 de la Clasificación Internacional de Enfermedades, CIE-9), incluidos pólipo y papiloma (código 236.7), en 14 hospitales de 4 Comunidades Autónomas de España. Cada caso nuevo diagnosticado entre el 1 de enero de 1985 y el 31 de marzo de 1986 que cumpliera los criterios de inclusión (ambos sexos, de menos de 80 años y residente en la provincia de localización de cada hospital) pasaba a formar parte del estudio. Todos los casos debían tener confirmación histológica según criterios pre-establecidos.

Por cada caso se incluyó un control ingresado en el hospital. El control hospitalario se seleccionó del libro de registro de admisiones y altas del mismo hospital que el caso. Se seleccionaba como control el primer individuo dado de alta por cualquier servicio (excluidos los de cirugía cardíaca y torácica) a partir de la fecha de diagnóstico del caso, que cumpliera los criterios de inclusión.

2. Selección de casos

Generalmente son personas que buscan atención médica y se les diagnostica el problema en estudio. Más raramente proceden de programas de diagnóstico precoz de enfermedades.

Un requisito importante es que sean *casos nuevos*, es decir, **casos incidentes**. Esto es esencial para evitar la mezcla de casos nuevos y antiguos que se produce si se usan *casos prevalentes*, como los procedentes de un corte transversal; si la enfermedad está relacionada con un grado importante de letalidad, por ejemplo, podrían estar sobre-representados los casos que por alguna razón han logrado sobrevivir. Entonces podría encontrarse una mayor proporción de esa característica en los casos que en los controles y deducirse que la exposición en estudio es un factor de riesgo.

Es bien conocido el ejemplo, referenciado por Breslow y Day ¹¹, de la aparente relación encontrada entre un tipo de antígeno HLA (HLA-A2) y el riesgo de desarrollar leucemia ⁴³. Se usaron como casos todos los diagnosticados de esa leucemia en el momento del estudio. Posteriormente, el seguimiento de los enfermos reveló que aquéllos con ese antígeno tenían un mejor pronóstico ⁴⁴. Por esa razón pareció, en un primer momento, que el HLA estaba asociado a un mayor riesgo de padecer leucemia.

La selección de los casos tiene como objetivo incluir todos los casos incidentes detectados en una *población bien definida* durante un *periodo de tiempo*. Cuando es difícil definir esa población (por ejemplo, si los casos son los ingresados en un determinado hospital) éso afectará a la forma de selección del grupo de controles. De cualquier manera suele ser necesario establecer un mecanismo de identificación de todos los casos que a veces es complejo.

Se necesita un **criterio diagnóstico** específico que delimite los criterios de inclusión junto a la definición de caso. Un criterio amplio, cáncer uro-genital por ejemplo, no serviría por cuanto incluiría tipos muy diferentes de neoplasias, asociadas lógicamente a factores de riesgo muy distintos. En esa situación, si alguno de esos cánceres está relacionado con una exposición determinada, se produciría una *dilución* de la asociación por la mezcla de casos y el estudio no podría identificarla.

El establecimiento de criterios de inclusión se complementa con la existencia de **criterios de exclusión**. Por ejemplo, es posible que se excluya a las personas más ancianas porque la realización de una encuesta sobre sus antecedentes de exposición sería poco fiable. Lo único que hay que mantener en mente es que si se aplica un criterio de exclusión en los casos, también habrá que aplicarlo a los controles.

3. Selección de controles

Posiblemente éste es el punto más difícil en el diseño. Los controles deben provenir de la misma *población* de la que proceden los casos. El objetivo es que proporcionen una estimación de la frecuencia de exposición que se observaría en los casos si no existiera asociación entre el factor en estudio y el efecto¹⁰. Una regla útil para decidir si un control es adecuado es que, si hubiera desarrollado el efecto, esa persona debería de haber aparecido en el grupo de casos.

Generalmente los controles son una muestra probabilística de la población donde se originan los casos. En ocasiones se prefiere elegir vecinos o amigos de los casos, aunque ésto tiene sus limitaciones. Cuando el origen de los casos está en hospitales o determinados centros asistenciales, entonces es más recomendable que los controles sean personas que acuden por otras causas a los mismos centros.

Lilienfield² propuso un esquema para la selección del grupo control según el origen de la serie de casos (Fig. I.44a-44f). Aunque no es la forma en que finalmente consideraremos la función del grupo control, su interés está en la introducción de una idea central: la comparabilidad de casos y controles a partir de su procedencia de una misma población, la que origina los casos.

La principal objeción al siguiente esquema es la ausencia de una referencia que más adelante introduciremos: que los casos sean casos incidentes aparecidos durante un periodo explícito de tiempo, en una población determinada (la población en estudio), o a partir de una experiencia de personas-tiempo (la base de estudio).

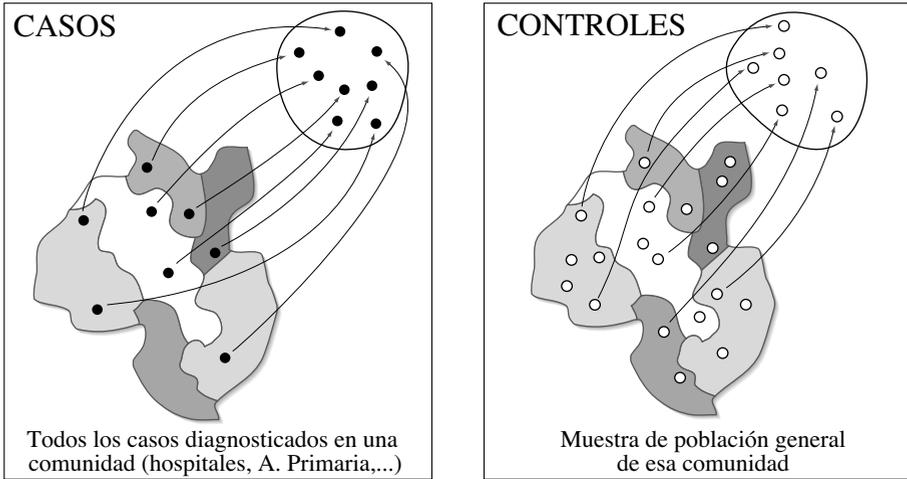


Fig. I.44a- Obtención de controles según la procedencia de los casos.

Si se puede garantizar la máxima exhaustividad en la localización de los casos, por ejemplo, si existe un Registro de Cáncer poblacional en la zona, entonces la población general representa bien el origen de los casos (Fig. I.44a). Se supone que son casos aparecidos durante un periodo de tiempo, no casos prevalentes.

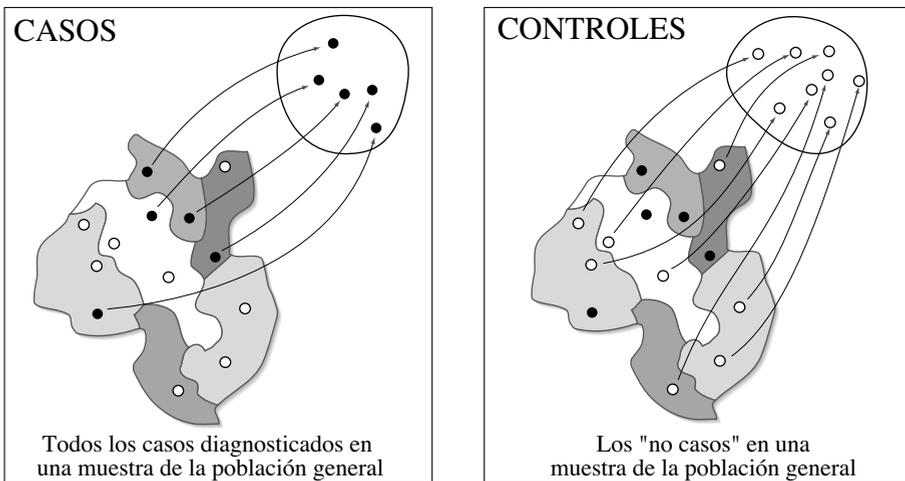


Fig. I.44b- Obtención de controles según la procedencia de los casos.

En este caso (Fig. I.44b) debe existir precaución, ya que si la muestra poblacional es un corte de tipo transversal, se tratará de *casos prevalentes*. Si es así podrían contener a casos supervivientes, y los factores relacionados con un mejor pronóstico aparecer como factores de riesgo, al ser más frecuentes en el grupo de casos.

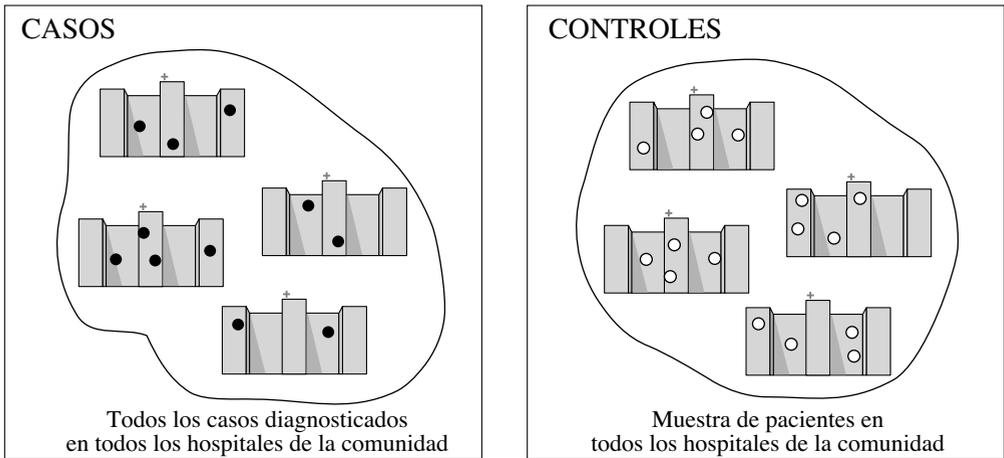


Fig. I.44c- Obtención de controles según la procedencia de los casos.

Como la probabilidad de que un caso sea ingresado depende de factores como la gravedad, accesibilidad, área real de influencia del hospital, entonces una muestra probabilística de la población no sería un buen grupo control (Fig. I.44c). Es posible que algunas personas en la población teóricamente adscrita al hospital, si desarrollaran el efecto, acudieran a hospitales de otro área, fueran tratadas de forma ambulatoria, etc.

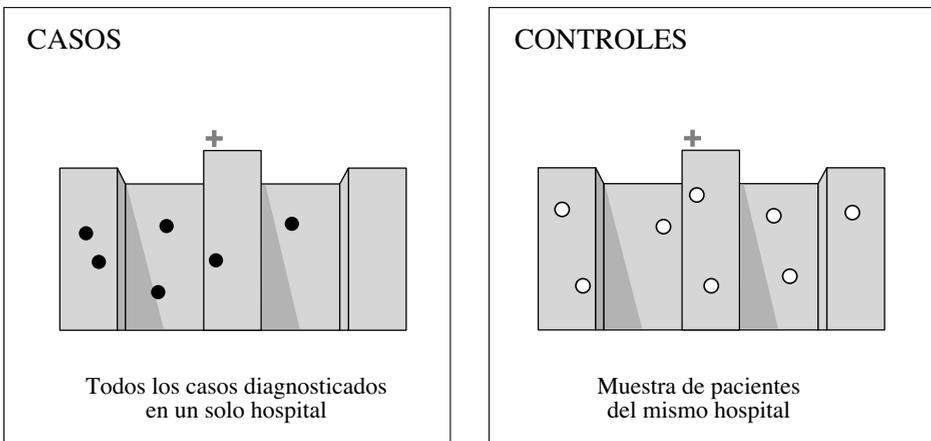


Fig. I.44d- Obtención de controles según la procedencia de los casos.

Por las mismas razones que en el caso anterior, es difícil en este caso conocer la población que origina los casos (Fig. I.44d). Se toma como grupo control una muestra de los individuos que habrían ingresado si hubieran desarrollado el efecto.

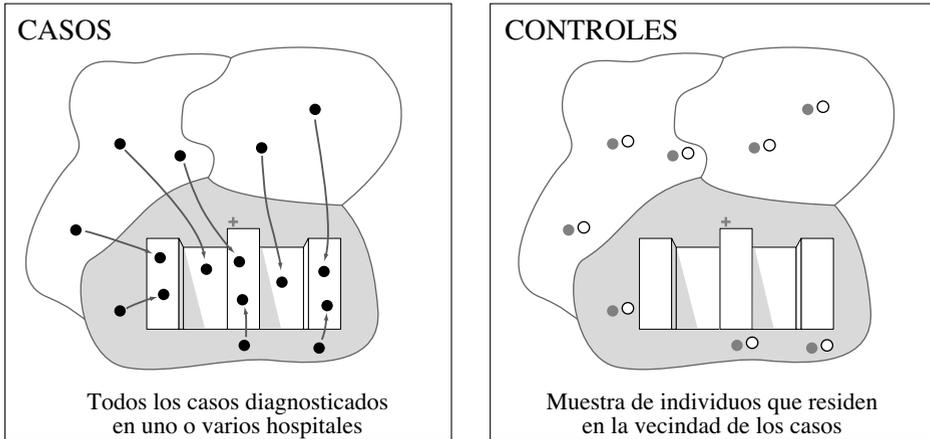


Fig. I.44e- Obtención de controles según la procedencia de los casos.

Los controles de vecindad tienen la ventaja de una mayor aceptación a participar (Fig. I.44e). Como es difícil definir la población que origina los casos, se supone que los controles de este tipo tendrían un destino semejante al de sus vecinos si se presentara el efecto. Sin embargo, se introduce un problema, como es el de que controles probablemente similares en situación socioeconómica, nivel cultural, exposiciones ambientales y hábitos de vida pueden ser muy semejantes a los casos incluso en la exposición en estudio.

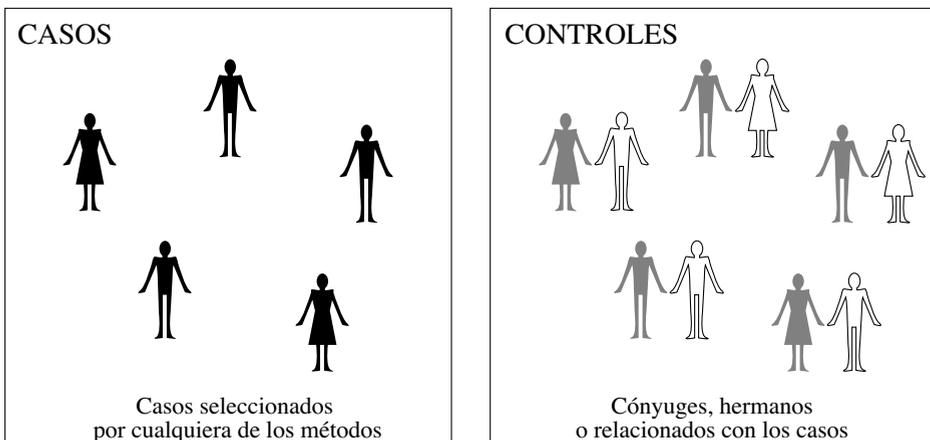


Fig. I.44f- Obtención de controles según la procedencia de los casos.

La selección aleatoria de convivientes (Fig. I.44f) presenta limitaciones que son semejantes al caso anterior, y es fácil que al hacer semejantes a casos y controles en una serie de variables, indirectamente se les pueda igualar en el factor que se estudia.

Estos ejemplos dan una idea, en principio intuitiva, y aplicada al concepto tradicional de los estudios caso-control, de la forma en que se deben seleccionar los controles más adecuados. Algo en común a todos ellos, y ésta es la idea principal que se pretende presentar, es la búsqueda de un grupo control representativo de la población origen de los casos. Un error en este punto podría distorsionar las estimaciones en la distribución de la exposición y llevar a conclusiones erróneas.

4. Grupo control: población en estudio y base de estudio

Según el concepto actual, cualquier estudio de casos y controles recoge durante un periodo de estudio los nuevos casos, *casos incidentes* (se excluye aquí al estudio caso-control que utilice casos prevalentes). Esos casos proceden de una población, la *población en estudio*⁴⁵, y han aparecido a lo largo de un periodo de tiempo, por lo que puede hablarse de una *cohorte* que los originó. Es decir, conceptualmente, cualquier diseño caso-control tendría un posible estudio de seguimiento equivalente, lo que se conoce como la *cohorte subyacente*. El problema es que el diseño de tipo cohorte encuentra frecuentemente incidencias muy bajas, lo que obliga a tamaños muestrales demasiado grandes para garantizar un número suficiente de casos, y esto puede ser poco eficiente. En su lugar, la estrategia del diseño de casos y controles busca *muestrear* la distribución de la exposición en la población que durante el periodo de estudio originó los casos⁴⁶. En definitiva, el muestreo consigue reducir los recursos necesarios (Fig. I.45).

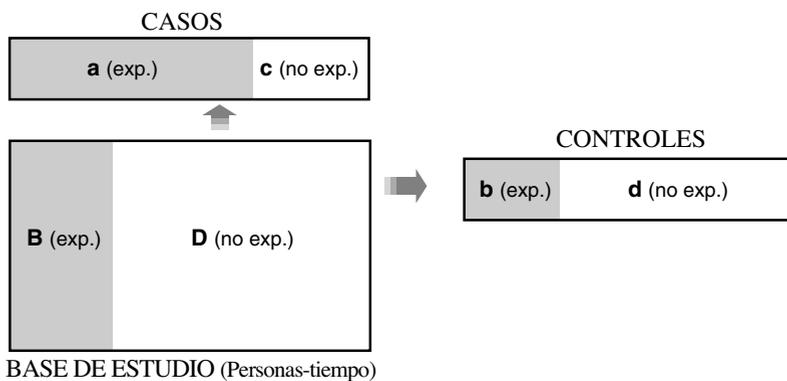


Fig. I.45- Casos y controles: población en estudio y base del estudio. Los controles son una muestra de la población o la base de estudio que origina los casos.

Modificado de Norell S⁴⁴.

La *población en estudio* de la que proceden los casos puede ser **cerrada**, es decir, una población fija en la que no se producen entradas ni salidas. Un estudio de seguimiento podría calcular incidencia acumulada si usara como denominador la población inicialmente en riesgo de que ocurriera la enfermedad. Puede calcularse el riesgo relativo (razón de riesgos) y la *odds ratio* de incidencia. En el estudio de casos y controles equivalente, los controles son representativos de la población de la que proceden los casos. Ofrecen por lo tanto información sobre la distribución de la exposición y la no exposición en la población si los controles se muestrean de la población al comienzo del seguimiento.

Si la *población en estudio* es **abierta** la única posibilidad en el estudio de seguimiento correspondiente es el cálculo de la *base del estudio* ('study base'), lo que se conoce como personas-tiempo en riesgo, en la cohorte expuesta y en la de referencia. Este análisis también es posible en cohortes cerradas. El diseño de un caso-control equivalente busca que los controles representen la distribución relativa de la exposición en el total de personas-tiempo que constituye la base del estudio. En este caso los controles se muestrean aleatoriamente, o conforme aparecen los casos (muestreo por densidad). Se espera que ese muestreo, que informa sobre la proporción de los expuestos y no expuestos, proporcione una estimación de la relación que existe entre el tamaño de la base de estudio en los expuestos en relación al de la base de estudio de los no expuestos.

En ambos tipos de poblaciones, abiertas y cerradas, como el interés está en conocer esa distribución en *toda* la población, es posible que el muestreo incluya casos en el grupo *control*. Si el grupo control se compone exclusivamente de *no casos*, es decir, personas sin el efecto, entonces las estimaciones que se obtengan se aproximan a las del estudio de seguimiento equivalente (la razón de riesgos) sólo si la enfermedad es poco frecuente. Si, por el contrario, el muestreo aleatorio de la población en estudio o de la base de estudio no excluye a casos en el grupo control, entonces se estima directamente la razón de riesgos o la razón de tasas sin necesidad de asumir que la enfermedad o el efecto es poco frecuente. Aunque la discusión sobre ese supuesto escapa a los objetivos de esta monografía ⁴⁷, y en adelante no distinguiremos entre ambas situaciones.

5. Recogida de la información

Los antecedentes de exposición frecuentemente son difíciles de medir con exactitud. Puede usarse entrevista con un cuestionario; en ese caso la duda es si el esfuerzo por recordar es igual de intenso en casos que en controles. Si el problema en estudio es grave, probablemente los casos recuerden mejor.

En ocasiones la información se puede obtener de registros disponibles como historias clínicas, en estudios caso-control de base hospitalaria. Esto evita en parte los problemas de recuerdo, aunque depende de la calidad de la información. En algunas situaciones se dispone de muestras biológicas, por ejemplo si se inició un estudio de seguimiento y se almacenaron muestras de sangre.

No toda la exposición se tiene en cuenta siempre. Puede ‘borrarse’ parte de ella, por ejemplo el año previo al diagnóstico, para garantizar que precedió a la aparición del efecto. Esto tiene interés sobre todo en procesos que pueden cursar con una fase preclínica prolongada, como algunos tipos de cáncer.

6. Sesgos

Los errores sistemáticos, especialmente los de mala clasificación, son generalmente más probables en un estudio de casos y controles que en su estudio de seguimiento equivalente.

6.1. Sesgo de selección

Afecta en la medida en que los sujetos participantes difieran de la población de la que proceden. Una situación frecuente es cuando los casos son personas ingresadas por su proceso en hospitales. La población de la que proceden es difícil de definir, ya que debe cumplirse la posibilidad de que cualquiera en esa población, si hubiera desarrollado el problema, habría aparecido en el grupo de casos. Si se piensa en el área geográfica que teóricamente atiende el hospital, puede ser que sólo una proporción de los diagnósticos se hagan en ese hospital, o que parte de la población acuda preferentemente a otro hospital o a otro área. Así, es difícil determinar quiénes hubieran acudido a ese hospital de haberse presentado el efecto. Las características que determinan el verdadero área de influencia son en general desconocidas, y podrían estar relacionadas con la exposición en estudio.

En ese caso obtener una muestra de población general como grupo control puede ser arriesgado. Si los casos y los controles representan a poblaciones distintas, se produce lo que se conoce como **sesgo de Berkson**⁴⁸. Para evitarlo, los controles pueden seleccionarse de entre personas ingresadas durante el mismo periodo en el hospital por otras causas, por ejemplo de distintos servicios, excluyendo diagnósticos relacionados positiva o negativamente con el factor de riesgo en estudio; para conseguir que el grupo control ofrezca una buena estimación de la distribución de la exposición en la población que originó los casos, las enfermedades incluidas como controles deben

tener una probabilidad de hospitalización semejante a la de los casos. Para evitar mayores problemas es recomendable además que el grupo control contenga una mezcla de diagnósticos distintos. Así una selección *distorsionada* de los casos se compensaría con una selección *distorsionada en el mismo sentido* de los controles, mejorando su comparabilidad. Se trata entonces de un estudio de casos y controles de base hospitalaria. Una ventaja añadida es que tanto casos como controles recuerdan con igual intensidad sus antecedentes de exposición, y la participación es más alta.

Aunque también son posibles, las **negativas** a participar son más frecuentes normalmente en los controles. Si los no participantes son sistemáticamente distintos en algo a los que participan, y si ese algo está relacionado con la exposición o el efecto, el grupo estudiado puede no representar adecuadamente a la población de la que procede. En ocasiones la única forma de evidenciar si se ha introducido algún tipo de sesgo es comparar respondientes y no respondientes en algunas variables de las que se pueda disponer información (sexo, edad, etc).

6.2. Sesgo de clasificación

Si la recogida de información está influenciada por el hecho de que alguien sea caso o control, se introduce inevitablemente una distorsión en las estimaciones. La posibilidad más frecuente es encontrar casos más expuestos y controles menos expuestos de lo que realmente están. Un ejemplo clásico es el de las madres de niños nacidos con alguna malformación; probablemente recuerdan mejor exposiciones durante el embarazo que las madres de niños sanos, que tienden a olvidar más. Si esto se produce, el error va en el sentido de exagerar diferencias, existan o no en la realidad. Se trata de una mala clasificación diferencial porque el error se produce más en uno de los grupos, el de los controles. Esto se puede producir por fallos de memoria, o por quienes recogen la información. Es recomendable que si van a participar encuestadores, se evite que conozcan quién es caso y quién es control. Como ésto es a veces inevitable, se debería impedir al menos que conozcan la hipótesis principal del estudio. Si la mala clasificación se produce de forma semejante en casos y controles, tiende a diluir la posible asociación. Es un error de tipo no diferencial. Una solución parcial para averiguar si existe mala clasificación de la exposición y en qué sentido influye es el uso de preguntas de control neutras: se refieren a variables para las que no se debería encontrar diferencias en casos y controles en ausencia de mala clasificación.

6.3. Sesgo de confusión: el apareamiento

Es posible que una variable de confusión se distribuya desigualmente en casos y controles. En muchas ocasiones se ha intentado igualar ambos grupos en relación a esa tercera variable apareando a los casos con los controles. Para cada caso, por ejem-

plo, se elige un control que sea semejante en cuanto a la variable que interesa homogeneizar: la misma edad, sexo, barrio, etc.

Sin embargo, el apareamiento puede introducir efectos indeseables: si se trata realmente de una variable confundente, debe estar asociada tanto con la exposición en estudio como con el efecto. Entonces, al hacer semejantes a ambos grupos respecto a la variable confundente, se les hace más semejantes también en cuanto a la variable en estudio. No ocurriría lo mismo si se apareara en un estudio de seguimiento, puesto que en ese caso se parte de cohortes expuestas y no expuestas. El desarrollo de las actuales técnicas multivariantes para el control estadístico de la confusión, en la fase de análisis, hace además menos interesante el uso del apareamiento. Por otro lado, si por alguna razón se decide aparear, el análisis estadístico debe tenerlo en cuenta, y no están indicados los métodos que se describen a continuación.¹¹

7. Análisis de la información

7.1. Magnitud o Fuerza de la asociación

Los estudios de casos y controles en los que se consigue información sobre todos los casos incidentes en un periodo de tiempo, y en los que se conoce la *fracción de muestreo* de los controles, permiten estimar la incidencia acumulada⁴⁹. Sin embargo ésa es una situación poco frecuente, y en general no es posible calcular incidencia, ya que no se conoce cómo se distribuye la aparición de casos en expuestos y no expuestos en la población en estudio. Por ello, no es posible el cálculo directo del Riesgo Relativo (*RR*). En su lugar éste se *estima* mediante la **Odds Ratio** de exposición (**OR**), traducida como **Razón de Ventajas** o **Razón de Oportunidades**.

Una primera aproximación intuitiva al sentido de la *OR* es la siguiente. Supongamos un factor de riesgo (*F*) que aumenta la probabilidad de que aparezca el efecto (*E*). Si es así, se encontrará una mayor proporción de antecedentes de exposición en relación a la no exposición en los casos que en los controles (Tabla I.6).

En el grupo de casos puede obtenerse un cociente entre la proporción de los que presentan antecedentes de exposición y la proporción de los que no tienen esos antecedentes. Este cociente relaciona dos probabilidades: la de que los casos presenten exposición y su complementaria, la de que no la presenten. Se conoce como **Odds** de *exposición* en los casos ($Odds_F(E)$, en la tabla). La palabra tiene su origen en las apuestas de las carreras de caballos (la *odds* o *ventaja* de un caballo en las apuestas). Se han propuesto traducciones como **Ventaja** u **Oportunidad**, pero no hay acuerdo general⁵⁰. En definitiva, es el cociente de una probabilidad por su complementaria.

	<i>E</i>	<i>no E</i>
<i>F</i>	<i>a</i>	<i>b</i>
<i>no F</i>	<i>c</i>	<i>d</i>

$$Odds_F(E) = \frac{\frac{a}{(a+c)}}{\frac{c}{(a+c)}} = \frac{a}{c} \quad Odds_F(no E) = \frac{\frac{b}{(b+d)}}{\frac{d}{(b+d)}} = \frac{b}{d}$$

$$O.R. = \frac{Odds_F(F)}{Odds_F(noF)} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a*d}{b*c}$$

Tabla 1.6- Odds Ratio de exposición en casos y controles. Cociente entre la odds de exposición de los casos y la odds de exposición de los controles.

De la misma forma puede calcularse la **Odds** de *exposición* en los controles (expresada como $Odds_F(no E)$ en la tabla). Si la hipótesis es cierta, y la exposición *F* se comporta como un factor de riesgo, entonces se espera que la odds de exposición sea mayor en los casos que en los controles. Pueden dividirse ambas *oportunidades*, y el cociente resultante, la **Odds Ratio** (Razón de Ventajas, Razón de Oportunidades,.. por resumir, la **OR**) será mayor que 1.

Si no hubiera asociación, no hay razón para esperar que la exposición aparezca de forma distinta en casos que en controles, y la **OR** será igual a la unidad. Si, por el contrario, la exposición tiene un efecto protector, entonces la **OR** será inferior a la unidad.

Como se vé, el comportamiento de la **OR** es idéntico al del riesgo relativo (**RR**). Pero ¿existe alguna relación entre ellos? Para solventar ésto, puede ser útil recordar la función del grupo control: proporcionar una estimación de la razón exposición/no exposición en la población de la que proceden los casos.

También debe recordarse que el Riesgo Relativo se define como la razón entre incidencia en expuestos e incidencia en no expuestos ($RR=I_F/I_{noF}$), y éso es así tanto si se usa incidencia acumulada como densidad de incidencia (Tabla I.7).

Evaluaremos la relación entre los estimadores de magnitud de asociación para el caso particular de una cohorte cerrada equivalente en la que podría calcularse el riesgo relativo a partir de la incidencia acumulada. La demostración para la *OR* de incidencia y la razón de tasas es similar.

	<i>E</i>	<i>no E</i>
<i>F</i>	<i>a</i>	<i>b</i>
<i>no F</i>	<i>c</i>	<i>d</i>

$$I_F = \frac{a}{(a+b)} \quad I_{noF} = \frac{c}{(c+d)}$$

$$RR = \frac{I_F}{I_{noF}} = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}}$$

Tabla I.7- Tabla 2x2 para calcular el riesgo relativo (razón de riesgos o RR) en estudios de seguimiento.

En el caso de la incidencia acumulada I_F puede expresarse como $I_F=a/(a+b)$ donde $(a+b)$ es la población en estudio en la cohorte expuesta. De la misma forma, $I_{noF}=c/(c+d)$ donde $(c+d)$ representa a la cohorte no expuesta. La expresión para el riesgo relativo puede reescribirse algebraicamente como:

$$RR = \frac{I_F}{I_{noF}} = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}} = \frac{a*(c+d)}{c*(a+b)} = \frac{\frac{a}{c}}{\frac{(a+b)}{(c+d)}}$$

En el estudio caso-control equivalente, a y c son conocidos (casos expuestos y no expuestos) (Tabla I.8).

	E	$no E$
F	a	b
$no F$	c	d

$$O.R. = \frac{Odds_F(E)}{Odds_F(noE)} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a*d}{b*c}$$

TABLA I.8- Tabla 2x2 para el cálculo de la odds ratio u OR (estudios de casos y controles)

El cociente $(a+b)/(c+d)$, es decir, la relación entre la población en estudio expuesta y no expuesta (F/noF) se estima (\simeq) a partir de la relación controles expuestos/controles no expuestos (b/d), cuando el grupo control es una muestra representativa de la población al comienzo del seguimiento. Por ello, puede hacerse la siguiente aproximación:

$$RR = \frac{\frac{a}{c}}{\frac{(a+b)}{(c+d)}} \simeq \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a*d}{b*c} = O.R.$$

Así, el riesgo relativo (RR) se estima (\simeq) como $(a*d)/(b*c)$, es decir, la OR . Y se estima porque las cantidades b y d proceden de un muestreo probabilístico, el que se usa para conseguir el grupo de controles. En ese caso, la interpretación de la OR es idéntica a la del Riesgo Relativo en términos de magnitud de la asociación (incrementos en el riesgo de padecer una enfermedad si está presente una determinada exposición, por ejemplo).

La OR es también un buen estimador de la razón de tasas, es decir, el riesgo relativo cuando se usa densidad de incidencia⁵¹. En ese caso, la razón controles expuestos/controles no expuestos (b/d) es una aproximación a la razón personas-tiempo expuestas/personas-tiempo no expuestas. Lo que es lo mismo, la razón entre la base de estudio en expuestos y la base de estudio en no expuestos⁵².

7.2. Medidas de impacto potencial

Si no es posible calcular incidencia, no se puede calcular directamente el **Riesgo Atribuible** en estudios de casos y controles. Sin embargo, otras medidas, como la **Proporción Atribuible** en expuestos y la **Proporción Atribuible Poblacional** se pueden obtener en diseños de casos y controles. Como se vió en los estudios de seguimiento, ambas estimaciones de impacto potencial se pueden calcular si se conoce el valor del Riesgo Relativo (que se estima en un caso-control mediante la Odds Ratio), y en el caso de la Proporción Atribuible Poblacional, la prevalencia del factor de exposición. Esto último es posible conocerlo en un caso-control si el grupo control es realmente representativo de la población de la que proceden los casos: la proporción de controles expuestos da idea de cuál es la prevalencia del factor en esa población.

Es una creencia extendida que sólo los diseños experimentales o cuasi-experimentales están indicados cuando se intenta evaluar intervenciones de cualquier tipo, sean éstas de tipo preventivo, curativo, etc. Sin embargo debe considerarse que cualquier intervención (como tipo de cirugía, programas de promoción de hábitos saludables o cuestiones relacionadas con la gestión de un hospital, por ejemplo) supone en definitiva someter a conjuntos de personas a una determinada exposición (F), relacionada con un efecto (E), que en este caso se considera beneficioso.

Por ello la valoración de cualquier intervención trata, en definitiva, de evaluar hipótesis variadas acerca de la asociación entre una exposición y un determinado efecto. En consecuencia, los diferentes diseños analizados hasta aquí pueden, con una mayor o menor potencia, utilizarse con ese propósito.

Por ejemplo, un estudio de casos y controles evaluó el impacto de un programa de detección precoz de cáncer de próstata⁵³. Se compararon 139 hombres con cáncer de próstata metastásico (estado D) con un número igual de hombres sin esa condición, todos afiliados a un seguro privado. La información se recogió de las historias clínicas. En los 10 años previos (excluyendo los 3 últimos meses) no se encontraron diferencias estadísticamente significativas entre los casos y los controles, en cuanto al número de exámenes rectales, por lo que no pudo concluirse que ese tipo de medida preventiva tuviera un impacto evidente.

Un estudio en la región de Florencia (Italia) comparó los antecedentes de contacto con un programa que ofrecía mamografías cada 2 años, en 57 mujeres muertas por cáncer de mama (casos) durante un periodo determinado, y 114 mujeres vivas residentes en el mismo área (controles)⁵⁴. Se encontró que el riesgo de morir por cáncer de mama, en el periodo de estudio, para las mujeres sometidas a screening frente a las que nunca habían acudido, bajaba a la mitad ($OR=0.57$).

Los diseños observacionales o no experimentales tienen por lo tanto un importante papel en lo que se conoce como *investigación evaluativa*, o evaluación de intervenciones. Más aún, cuando la evidencia procedente de este tipo de estudios es lo suficientemente fuerte, en ocasiones se toman decisiones sin esperar a una posterior evaluación experimental.

Por otro lado, se sabe que es recomendable el establecimiento de programas de intervención con el objetivo de disminuir la prevalencia de la dependencia del tabaco y otras drogas, aún con ciertas reservas⁵⁵. Los principales argumentos provienen de estudios no experimentales. Sería práctica y éticamente difícil, con las pruebas disponibles, defender una *demonstración final*, experimental, sobre esas evidencias. Además, factores relacionados con estilos de vida, dieta, etc, frecuentemente asociados con enfermedades de tipo crónico, son difícilmente abordables mediante experimentación. Puede entenderse este argumento mejor imaginando un diseño experimental (con asignación aleatoria) que intentara evaluar los potenciales efectos beneficiosos de la siesta sobre la calidad de vida; las dificultades para conseguir participantes en el experimento serían enormes sobre todo en el grupo control (no siesta).

Lo único que diferencia a los diseños no experimentales de los experimentales en este sentido es que los primeros no incorporan la *intervención* en su propio diseño. Y, por supuesto, no existe la posibilidad de manipular variables (intervenir) o decidir respecto a quién se expone o no; éso ocurre independientemente de la voluntad de los investigadores.

G. ESTUDIOS CUASIEXPERIMENTALES Y ESTUDIOS ALEATORIOS CONTROLADOS

E. Perea-Milla López

1. Introducción

Se engloban aquí un conjunto de diseños en los que existe una intervención planificada por los investigadores y hay algún control sobre la asignación de los participantes a los grupos de comparación. Aunque, como veremos después, no siempre están presentes ambos aspectos.

La **intervención** consiste en la manipulación de una variable; se somete a un grupo a alguna exposición y se compara con un grupo no expuesto que se usa como referencia o control. La variable de exposición es la que se manipula y el objetivo es medir su efecto⁵⁶. Como es lógico la exposición consciente de seres humanos sólo es posible si se presume que tendrá efectos beneficiosos para ellos. Por eso este tipo de diseños tiene restringido su uso a evaluación de intervenciones preventivas, nuevos tratamientos o nuevas tecnologías. Los estudios experimentales son los únicos en los que la intervención está integrada en el diseño. Lo que no excluye la posibilidad de evaluar impacto de intervenciones usando aproximaciones no experimentales.

Por **control de la asignación** se entiende que al formar los grupos de comparación, el investigador decide en cuál estará incluida cada persona que participe. Como es obvio, esa decisión no debe ser caprichosa. En los experimentos puros se usa la asignación aleatoria, es decir, el azar. En los estudios *menos* experimentales la asignación aleatoria *individual* no es posible y si hay alguna participación del azar, afectará a grupos enteros y no a individuos. Esto último define a los diseños **cuasiexperimentales**.

En el caso de los **estudios aleatorios controlados** o **ensayos clínicos**, la asignación de los participantes es aleatoria individual. Aquí el azar decide, para cada individuo, el grupo en el que va a participar. El objetivo es que, si el tamaño muestral es suficientemente grande, los grupos sean homogéneos y exista la máxima comparabilidad entre ellos. De esta forma, se espera que otras variables distintas a la de exposición y que pudieran interferir en los resultados, se distribuyan de forma semejante en los grupos de comparación. Así se neutraliza el efecto confundente que una distribución desigual podría introducir.

2. Estudios cuasiexperimentales

Podrían definirse como estudios experimentales en los que falta o bien la asignación aleatoria o bien el grupo control ⁵⁷. Veney ⁵⁸ propuso un esquema del planteamiento general de estos diseños (Fig. I.46). Se trata de responder a la pregunta: ¿conseguió la intervención producir alguna diferencia? Las intervenciones a evaluar pueden ser muy diversas. Por ejemplo, interesaría saber si el tipo de formación e información que se facilita a los pacientes en diálisis o transplantados, al alta, influye sobre su calidad de vida. O si el hecho de recibir periódicamente en un servicio información sobre el coste e indicaciones de las pruebas complementarias mejora la calidad y la eficiencia de su prescripción. En situaciones de ese tipo puede medirse algún indicador de la situación (lo que en cuasiexperimentación se conoce como *indicador de estado* del sistema). A continuación se introduce alguna modificación (la intervención), y una segunda observación de nuevo mide la situación. La diferencia entre la primera y la segunda observación determina si la modificación tuvo algún efecto.

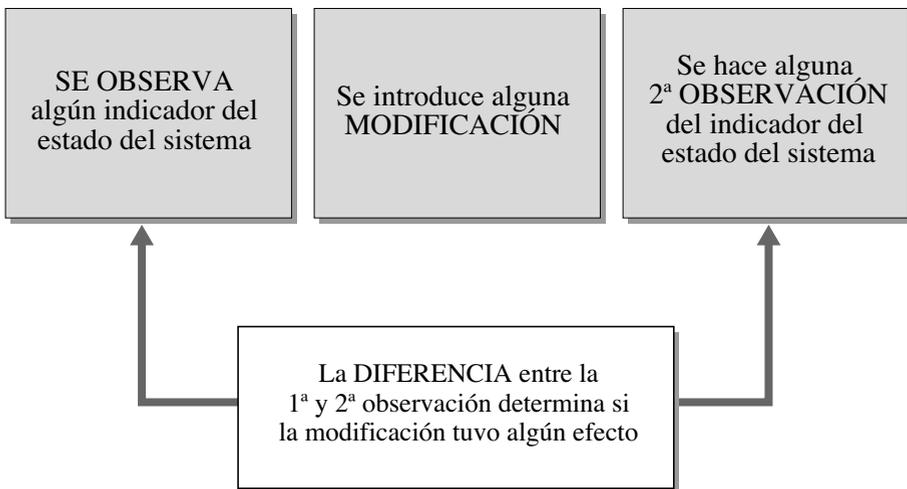


Fig. I.46- Esquema de un estudio cuasiexperimental. Tomado de Veney J. ⁵⁸

Este esquema puede complicarse. Por ejemplo, sería interesante saber si un grupo semejante a aquél sobre el que se interviene tiene una evolución similar en el tiempo a pesar de no sufrir la intervención. En ese caso se está usando un grupo no expuesto como grupo de referencia o *control*. Distintas combinaciones producen los diversos tipos de estudios cuasiexperimentales. Revisaremos sólo dos tipos, simples pero ilustrativos: los diseños **antes-después** (o **pretest-postest**), y los del tipo **antes-después con grupo control**.

2.1. Estudios antes-después (pretest-postest) sin grupo control

En este caso el único grupo considerado es el que se conoce como **grupo experimental** (Fig. I.47). En un primer momento se hace una observación, que se nota como O_1 . Puede tratarse de mortalidad postoperatoria, tasa de infecciones, morbilidad poblacional o cualquier otra medición que interese. Se obtiene así información sobre el nivel basal o de partida. A continuación se introduce el efecto experimental, notado como X ; puede ir desde un programa complejo, como el establecimiento de contratos-programa con los hospitales, a una actividad concreta, como la introducción de un programa de educación sobre autocuidados. En cualquier caso, tras la intervención se hace una segunda observación (O_2), que idealmente debería usar el mismo tipo de medición que la primera.

En realidad puede tratarse de una combinación de diseños no experimentales. Por ejemplo, un primer estudio transversal o de prevalencia, seguido de una intervención y de un segundo corte de prevalencia. Esto tendría sentido si la intervención tiene algún efecto sobre la incidencia de una enfermedad o sobre su pronóstico. En el caso de intervenciones de tipo preventivo, una reducción de la incidencia podría evidenciarse mediante un estudio de seguimiento, por ejemplo, de incidencia de infección nosocomial. El uso de análisis de series temporales, usando datos agregados, es de una enorme utilidad en este terreno (en este caso el nivel de análisis es ecológico, al comparar cambios en la tendencia a lo largo del tiempo). El diseño cuasiexperimental puede usar por lo tanto secuencias de estudios observacionales, incorporando la intervención al propio estudio.

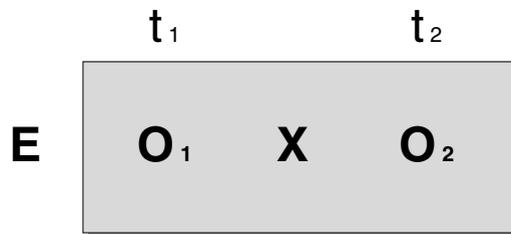


Fig. I.47- Estudio antes-después (pretest-postest) sin grupo control.

La comparación de las dos mediciones, O_1 y O_2 , dá una estimación del efecto de la intervención. Sin embargo, en la mayoría de los casos esa diferencia puede deberse a causas ajenas a la propia intervención, como se discute a continuación.

Por un lado, la influencia de **otras variables**, distintas de la intervención, podría explicar las diferencias encontradas. Puede imaginarse el caso de la introducción de una nueva técnica quirúrgica que disminuya el riesgo de infección postoperatoria. Si simultáneamente se hubiera instaurado una protocolización de la quimioprofilaxis preoperatoria, ¿cómo deslindar el impacto de cada una de esas intervenciones?. Un descenso en la tasa de infección podría estar influenciado por ambas, sólo por una de ellas, o por cualquier otro cambio que se hubiera producido paralelamente en el tiempo. Esa distorsión, el efecto de variables confundentes, no puede eliminarse con este diseño. Un problema relacionado es la comparabilidad de los grupos antes y después. Es frecuente en la literatura el uso de grupos de control **históricos**. En ese caso la preocupación máxima es conocer si, para distintas variables relevantes (como gravedad, situación general,..), los dos grupos que se comparan son similares.

Los participantes en el estudio también pueden contaminar los resultados. La **sensibilización** de los sujetos estudiados, conocedores de su participación en el experimento, por ejemplo. La percepción y comunicación de intensidad del dolor, la satisfacción de los usuarios con el hospital, son variables que pueden cambiar por el hecho de saberse observados y no son una consecuencia estricta de la intervención. Las expectativas del investigador y del personal relacionado con el estudio pueden influir también, y ésto, que de alguna forma es inevitable, es un problema si posteriormente se decide generalizar la intervención de forma *rutinaria*.

Si la variable que se mide presenta oscilaciones a lo largo del tiempo, por ejemplo la mortalidad postoperatoria, surge un nuevo problema. Es posible que un año, por mera variación aleatoria, la tasa de mortalidad esté por encima de la tendencia general. Parece sensato esperar que el año siguiente, si nada ha cambiado, regrese al nivel *habitual*. Si en medio se produjo una intervención, aparentemente ésta tuvo entonces un impacto. Sólo que si no se hubiera producido, la mortalidad igualmente hubiera descendido. Este fenómeno se conoce como **regresión a la media**, y un diseño antes-después no puede controlarlo. Se necesitaría conocer además el comportamiento de la mortalidad durante los años previos.

Aún con estas limitaciones, son frecuentes las ocasiones en las que no es posible otro tipo de diseño. Porque la intervención ya ha comenzado, porque no es posible encontrar un grupo comparable que sirva como control, etc. En ése caso lo importante es no perder de vista las posibles explicaciones alternativas a la que asegura que la intervención fue exitosa.

2.2. Estudios antes-después con grupo control

Utilizan un grupo *experimental* (E), en el que se produce la intervención, y otro grupo de *control* (C), sobre el que no se interviene (Fig. I.48). Este último se conoce como grupo control *no equivalente* (sin asignación aleatoria individual) en la literatura sobre cuasiexperimentación. La razón es que no hay asignación aleatoria individual. Podría elegirse al azar el grupo sobre el que se va a intervenir, pero desde el momento en el que éso se decide, todos sus integrantes serán expuestos. Por ejemplo, los enfermos correspondientes a una sección u otra dentro de un mismo servicio hospitalario. Como no hay asignación individual, no se puede garantizar la completa comparabilidad de los grupos.

En un primer momento se hace una observación en ambos grupos, el experimental (O_{E1}) y el control (O_{C1}). La intervención (X) tiene lugar sólo en el grupo experimental. Posteriormente se repite la medición, también en los dos grupos (O_{E2} y O_{C2}). Se espera que las diferencias encontradas no sean iguales, precisamente debido a la intervención.

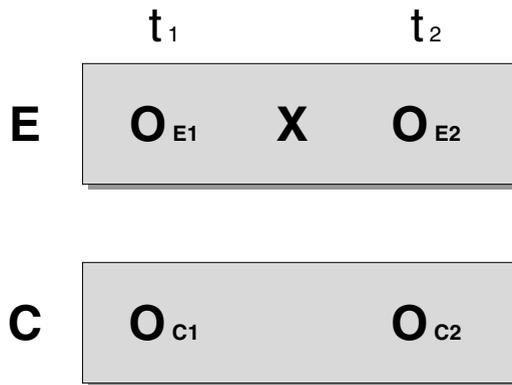


Fig. I.48- Estudio antes-después con grupo control.

Un diseño de este tipo podría llevarse a cabo con el presunto conjunto de lectores de esta monografía. Si antes de comenzar su lectura se pudiera evaluar el nivel de conocimientos previo (por ejemplo mediante preguntas tipo test), al final se podría repetir esa prueba o una semejante. Si ésto se hiciera con todos los lectores, el diseño sería lógicamente de *antes-después*. La diferencia de nivel estimaría la hipotética utilidad de la lectura. Pero también se podría atribuir a estar asistiendo simultáneamente a un curso de actualización, por ejemplo, y no se podría entonces evaluar el impacto aislado de la monografía.

¿Cómo hacerlo? Una forma, desde el punto de vista cuasiexperimental, sería decidir que aquéllos procedentes del hospital A leyeran la monografía en cuestión, y los del hospital B no lo hicieran. Tendríamos entonces un estudio *antes-después con grupo control*. El problema probablemente se centre entonces en la comparabilidad de ambos hospitales, habilidades y destrezas de los que proceden de cada uno, etc. Cuestiones que dado que se trata sólo de un ejemplo, es posible dejar de lado.

La presencia de un grupo control permite solucionar en parte algunos de los problemas discutidos antes. Si se puede asumir que la influencia de **otras variables** aparece por igual en los dos grupos, entonces puede considerarse controlado su efecto. Sin embargo es posible que ésto no sea cierto, y que los enfermos de una sección sean distintos a los de otra, en promedio, en algo que pudiera contaminar los resultados. Lo mismo ocurre con los médicos de varios hospitales o los habitantes de distintas zonas. Es el resultado de no poder asignar aleatoriamente de forma individual. Como en los estudios no experimentales, será necesario entonces reunir información sobre las variables potencialmente confundentes y tenerlas en cuenta en el análisis. Además, por supuesto, de buscar grupos que sean lo más parecidos (comparables) entre sí en la fase de diseño.

La **sensibilización** de los participantes es de esperar que se produzca tanto en el grupo experimental como en el control. De hecho, las observaciones se hacen de la misma forma en los dos grupos, tanto antes como después. Entonces, si las diferencias son mayores en el grupo experimental, podría aislarse el efecto de la intervención del atribuible a la sensibilización. Las expectativas de los investigadores y del personal que participan, sin embargo, pueden aquí influir, ya que se está en condiciones experimentales y no rutinarias.

Si se puede asumir que los grupos de comparación son semejantes en cuanto al problema en estudio o a factores relacionados, es esperable que el efecto de la **regresión a la media** se comporte de forma idéntica en ellos. Si embargo es difícil evaluar ésto si no se conoce la tendencia global del problema con datos previos.

3. Estudios aleatorios controlados o ensayos clínicos

Son el *experimento* en las condiciones de control más ideales (Fig. I.49). Se realizan mediante asignación aleatoria individual, de forma que el investigador decide la exposición o no de los participantes usando el azar como criterio. En su forma más simple, lanzando una moneda al aire; también pueden usarse tablas de números aleatorios publicadas, o técnicas de computación. Se espera que el azar distribuya a las personas de forma homogénea, consiguiendo grupos comparables, simi-

lares en todo excepto en la exposición de interés. La manipulación de una variable consigue entonces que en uno de los grupos esté presente la exposición y en otro, que se usa como referencia o control, no lo esté. De esta forma, las diferencias que se encuentren al final del estudio pueden atribuirse al efecto de la exposición.

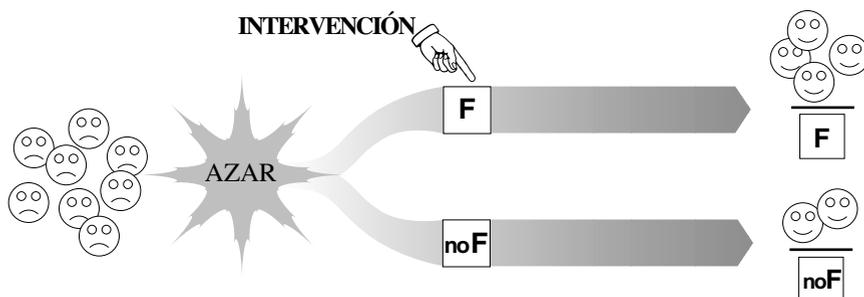


Fig. 1.49- Estudio aleatorio controlado.

La exposición debe ser presumiblemente beneficiosa para los individuos que la soportarán. Puede consistir en una actividad preventiva, un nuevo tratamiento, o una nueva tecnología. Para impedir daños producidos por una mala práctica en experimentación se ha desarrollado una legislación específica, que arranca con el Código de Nuremberg, después de la II Guerra Mundial. En España, los ensayos clínicos están regulados por la Ley del Medicamento (Ley 25/1990) y el Real Decreto sobre requisitos para la realización de ensayos clínicos con medicamentos⁵⁹ (Real Decreto 561/1993). Esta legislación se inspira en diversas Directivas de la Unión Europea⁶⁰.

3.1. Investigación clínica y preclínica

La evaluación de nuevos tratamientos, la mayoría de las veces medicamentos, se desarrolla en fases sucesivas. Incluye **estudios preclínicos**, que mediante experimentos de laboratorio, in vitro o en animales, estudian la farmacología y toxicología del nuevo medicamento. Más adelante, los **estudios clínicos** evalúan su comportamiento en humanos. Se comienza con una evaluación toxicológica inicial, tras la que es posible comparar la eficacia del nuevo tratamiento con uno estándar o con un placebo. Aquí es donde se sitúan los estudios aleatorios controlados. Una vez que el tratamiento está disponible de forma generalizada, queda sometido a procedimientos de *vigilancia terapéutica*.

De acuerdo con los objetivos perseguidos, la investigación de estudios clínicos se desarrolla en fases. Así, pueden distinguirse (Fig. I.50):

Estudios en fase I⁶¹. Son el primer paso de la investigación de una sustancia o medicamento nuevo en el hombre. Consisten en una evaluación de su farmacocinética y farmacodinámica (metabolismo, biodisponibilidad), que proporciona información preliminar sobre la *seguridad* del producto en términos de tolerancia, efectos adversos, etc. Se llevan a cabo en sujetos sanos o en algunos casos de pacientes, y orientan sobre la pauta de administración más apropiada para estudios posteriores.

Estudios en fase II⁶². Se realizan en pacientes que padecen la enfermedad de interés. Proporcionan información preliminar sobre la *eficacia* del producto, establecen la relación dosis-respuesta y el método óptimo para administrar el tratamiento, y amplían los datos sobre seguridad obtenidos en la fase I.



Fig. I.50- Fases de los estudios clínicos.

Estudios en fase III⁶³. Son los estudios aleatorios controlados propiamente dichos. Evalúan la *eficacia relativa* y la seguridad del tratamiento, intentando reproducir las condiciones de uso habituales. Se compara la eficacia del nuevo tratamiento con las alternativas terapéuticas disponibles previamente para la indicación estudiada, o con un placebo si no las hubiera. Se llevan a cabo en una muestra de pacientes más amplia, representativa de la población de pacientes a la que iría destinado el medicamento.

Estudios en fase IV. Se hacen cuando un medicamento ha sido comercializado, después de la aprobación para uso general. El objetivo es investigar los efectos a largo plazo. Pueden diseñarse estudios de seguimiento no experimentales o usar el sistema

de detección de efectos adversos (farmacovigilancia). Si se estudia algún aspecto aún no valorado o condiciones de uso distintas de las autorizadas, como una nueva indicación, los estudios en esta fase serán similares a los de las fases I, II y III.

Aún cuando la evaluación de nuevos medicamentos es el objetivo de la mayoría de los ensayos clínicos que se publican, no deben olvidarse otras utilidades. Retomando ejemplos anteriores, podría aleatorizarse a los pacientes que van a recibir información sobre autocuidados y comparar una nueva sistemática con alguna ya preexistente. El servicio de admisión de un hospital puede asignar al azar la cama donde cada enfermo va a ser ingresado y evaluar cambios organizativos en alguno de los servicios o secciones. La eficacia de la lectura de esta monografía podría medirse haciendo que, al azar, la mitad de los asistentes a cursos de actualización la leyeran y la otra mitad no. Al utilizar asignación aleatoria individual, se espera que los procedentes de distintos centros, con formación previa desigual, etc, se distribuyeran homogéneamente. Aunque en el último ejemplo el tamaño muestral pudiera ser claramente insuficiente.

Aleatorización. La *asignación* aleatoria no debe confundirse con el *muestreo* al azar a partir de una población definida. El proceso de aleatorización trata de la localización de los participantes en los grupos de comparación. El muestreo probabilístico, sin embargo, consiste en la obtención de un grupo representativo de la población en estudio. La asignación puede ser aleatoria *simple*, en cuyo caso todos los participantes tienen la misma probabilidad de formar parte de cada grupo.

En ocasiones se usa la técnica de la aleatorización *balanceada* mediante bloques, para asegurar que el número de personas incluídas sea el mismo durante todo el estudio. Si se decide garantizar que alguna variable relevante tenga un número suficiente de sujetos en sus distintos niveles, puede acudirse a la asignación *estratificada*. Dentro de cada estrato se asigna a los individuos a los grupos de comparación.

Otra posibilidad es usar un diseño de tipo *cruzado* (Fig. I.51). En ese caso cada participante recibe ambos tratamientos, el experimental y el control. La forma más común es en dos períodos: durante el primero la mitad de los participantes están en el grupo experimental y la otra mitad en el control. A continuación se cruzan, y se invierten las exposiciones. De esta forma se requiere un menor número de participantes, aunque el seguimiento de cada uno es mayor. Y se produce un *autoajuste*, ya que cada persona sirve como su propio control.

El problema es que el resultado en estudio puede impedir el cruce (por muertes, contraindicaciones,...), y además existe la posibilidad de que el efecto del primer tratamiento recibido se superponga al del segundo, por lo que frecuentemente se nece-

sita un periodo de *limpieza* posterior al primer tratamiento, durante el que no se suministra tratamiento. Esta es una limitación que suele impedir el uso de este diseño. Sería de utilidad, por ejemplo, en la medicación de distintas crisis dolorosas por litiasis renal.

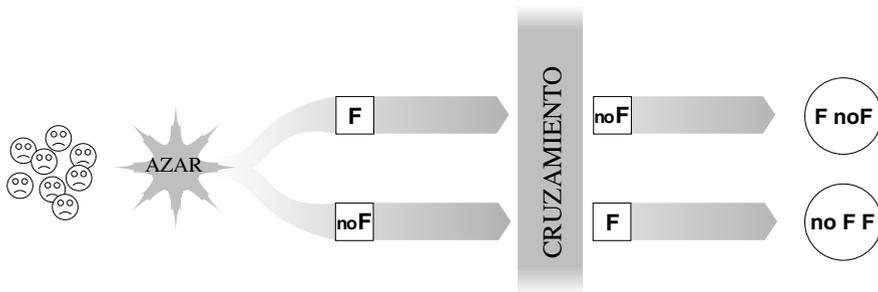


Fig. I.51- Diseños cruzados.

Los diseños de tipo *factorial*^{5,64} son una opción cuando se pretende verificar más de una hipótesis. Por ejemplo, el uso de analgésicos y espasmolíticos en el dolor cólico de la litiasis, y el papel de los ansiolíticos como coadyuvantes. Se podría aleatorizar primero y proporcionar analgésicos y espasmolíticos a un grupo y sólo analgésicos al grupo control (Fig. I.52).

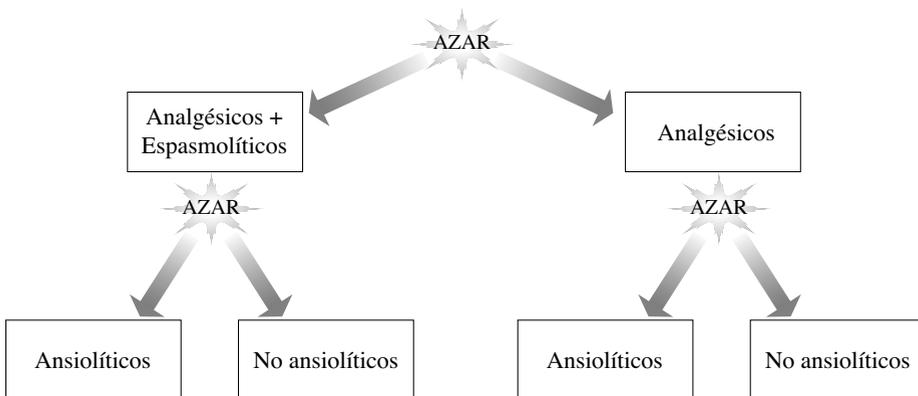


Fig. I.52- Diseño factorial.

A continuación se vuelve a aleatorizar y en cada grupo se define un subgrupo experimental, que recibe ansiolíticos, y otro de control que no los recibe. Al final del estudio se tienen cuatro grupos: los que reciben todos los tratamientos, sólo analgésicos y espasmolíticos, analgésicos y ansiolíticos, y sólo analgésicos. Las distintas comparaciones entre ellos permiten evaluar la eficacia independiente de cada tratamiento y sus posibles efectos sinérgicos y antagónicos.

3.2. Grupo control

Es el que se usa como referencia. Se espera que sea lo más similar posible, en cuanto a variables relevantes, al grupo experimental (en el que se produce la intervención que se evalúa). La distribución de las variables potencialmente confundentes sería entonces similar en los grupos de comparación.

Como se ha podido ver, este grupo no tiene porqué estar *no expuesto*. Su caracterización en ese sentido tiene que ver con el hecho de que *no está expuesto a la exposición en estudio* (un nuevo tratamiento, etc). Sin embargo, está muy extendida la creencia de que el grupo control es uno sometido a placebo. Pero si existe un tratamiento estándar previo, éste es el que se aplica en el grupo control. La idea es comparar el nuevo tratamiento a la mejor terapia disponible hasta ese momento. Si no existiera, tiene sentido usar entonces algún tratamiento *no activo* en el grupo control. En algunas ocasiones lo que se compara es dos nuevos tratamientos⁶⁵; en ese caso la designación de grupo experimental o control es arbitraria.

¿Cuándo usar entonces placebo? Se admite su utilización cuando no existe una terapia establecida para la enfermedad, la que existe no tiene una eficacia probada, o tiene una frecuencia elevada de efectos adversos. No debe olvidarse en este punto que el placebo puede ser eficaz, y puede ser útil sobre todo si además la enfermedad tiene poco impacto sobre la situación general del paciente. Se ha sugerido incluso que el conocimiento por el paciente del hecho de estar sometido a experimentación puede anular el efecto placebo que toda intervención médica comporta⁶⁶.

3.3. Sesgos

La asignación aleatoria no garantiza la inexistencia de errores sistemáticos; no hay que olvidar además que el azar puede jugar malas pasadas. Por lo demás, un estudio aleatorio controlado puede verse como un estudio de seguimiento o de cohortes, en el que se aleatoriza al comienzo.

a. Sesgo de selección

Los pacientes que se estudian han de ser una muestra representativa de los pacientes que presentan el problema en estudio, o al menos de aquellos a los que se dirige la indicación evaluada. Por ello, un primer problema lo presentan las personas que se niegan a participar en el estudio. El análisis debe contemplar si existen diferencias entre quienes participan y quienes no, para detectar posibles evidencias de sesgo de selección. Un problema semejante proviene de los abandonos durante el seguimiento.

La legislación española obliga al *consentimiento informado* (ver Facsímil), es decir, el sujeto debe expresar voluntariamente su intención de participar en el ensayo clínico, después de haber comprendido los objetivos del estudio, beneficios, incomodidades y riesgos previstos, alternativas posibles, derechos y responsabilidades. El documento de consentimiento informado debe figurar además en el protocolo del ensayo para que pueda ser aceptado por el Comité Ético de Investigación Clínica correspondiente.

Nombre del sujeto

Protocolo del Ensayo Clínico autorizado número

DECLARO QUE:

Se me ha solicitado participar en el Ensayo Clínico de referencia, durante el cual me será administrado un medicamento en investigación.

Antes de prestar mi consentimiento, firmando este documento, he sido informado por el investigador principal don..... del proyecto de investigación de que se trata, del tipo de medicación que va a serme administrado, de sus posibles efectos beneficiosos y también de sus inconvenientes, riesgos y reacciones adversas que pueden presentarse, de otras posibilidades existentes de tratamiento, de los propósitos del mencionado estudio y de la metodología con que va a ser llevado el mismo, así como de que por prestar mi colaboración libre y voluntariamente puedo suspenderla en cualquier momento que lo desee.

También se me ha informado de que para mi protección en el desarrollo del estudio dispondré de la asistencia médica y sanitaria adecuada y de los tratamientos complementarios que necesite y de que se tomará con el protocolo del ensayo todas las medidas apropiadas para que el estudio esté debidamente controlado.

(Firma del interesado)

(Lugar y fecha)

Recibí: El Investigador,

*Facsímil de la hoja de consentimiento informado.*⁵⁹

b. Sesgo de clasificación

La existencia de la aleatorización no elimina la posibilidad de una mala clasificación. La adherencia al tratamiento, por ejemplo, puede suponer distorsiones. El investigador no puede imponer lo que las personas deben hacer. Sólo puede alentar con firmeza a los individuos para que cumplan el plan terapéutico, y eviten medi-

camentos o procedimientos que pudieran interferir. Quizá en ocasiones no se puedan comparar tratamientos ‘puros’, sino sólo estrategias terapéuticas. El análisis no debería por lo tanto eliminar a los sujetos incumplidores (hecho frecuente en la literatura) ni a los que no aceptan participar.

Puede ocurrir que algunos pacientes cambien de grupo después de la aleatorización. En ese caso se está rompiendo con la regla de oro, ya que los que cambian pueden tener características en común e introducir un error de tipo sistemático. Si no se detecta, además, se les estará clasificando incorrectamente en cuanto a la exposición. Otra amenaza, la mala clasificación del efecto, debe controlarse consiguiendo mecanismos que garanticen que se buscan los resultados de la misma forma y con los mismos criterios en los grupos que se comparan.

Un buen mecanismo de control es el *enmascaramiento*; se trata de conseguir un cegamiento de los participantes para evitar la subjetividad en los resultados. En su forma más simple, el sujeto desconoce el grupo de tratamiento a que ha sido asignado; es el *simple ciego*. Si además tanto el sujeto como el investigador desconocen el grupo de tratamiento, se trata de un *doble ciego*. Cuando incluso el analista de los datos está cegado, se está realizando una evaluación con *triple ciego*.

3.4. Análisis

Como se ha dicho, los estudios aleatorios controlados pueden asimilarse a estudios de seguimiento o de cohortes, sólo que aquí se aleatoriza al comienzo. Por ello, la estimación de las medidas de magnitud y de impacto potencial son idénticas a las ya conocidas (capítulo I.E). Es frecuente encontrar en la literatura *análisis de supervivencia* en ensayos clínicos. La idea es que los sujetos en estudio no son seguidos todos necesariamente a partir de una misma fecha. Cada uno se sigue a partir de un punto bien definido, conocido como tiempo cero, y pueden ordenarse entonces los tiempos de seguimiento observados de menor a mayor. Se obtiene así una cohorte ordenada por tiempos, hipotética, que permite estimar la probabilidad de sobrevivir a distintos tiempos desde el punto cero (Capítulo II.D.5). Esa probabilidad se calcula a partir de la probabilidad acumulada de sobrevivir a cada uno de los intervalos de tiempo que les precedieron⁶⁷.

Los pacientes van introduciéndose en el estudio conforme pasa el tiempo. Sus respuestas al tratamiento se observan secuencialmente y se van acumulando. Es necesario por lo tanto un seguimiento adecuado, lo que se conoce como monitorización. Esto tiene interés para detectar desviaciones en el cumplimiento del protocolo y efectos adversos, y para el correcto procesamiento de los datos y el mantenimiento de

un sistema de información general. Pero además puede existir interés en ir haciendo *comparaciones intermedias* de tratamientos⁶⁸. Su utilidad está en el uso de los datos conforme van apareciendo para buscar diferencias entre tratamientos que pudieran ser convincentes para detener el ensayo. Se puede optar por una monitorización *continua*, aunque lo más frecuente es la monitorización *secuencial*. En este último caso se realizan pruebas de significación estadística repetidas a intervalos periódicos. Debe estar decidido a priori cuál es la diferencia suficientemente fuerte para detener el estudio. La parada se produce cuando se encuentra una diferencia de tratamiento estadísticamente significativa para alguna medida de resultado importante, en un nivel preacordado.

REFERENCIAS

1. Ahlbom A, Norell S. *Fundamentos de Epidemiología*. Madrid: Siglo XXI Ed., 1987.
2. Lilienfield A, Lilienfield D. *Fundamentos de Epidemiología*. México: Ed. Fondo Educativo Interamericano, 1983.
3. Kleinbaum D, Kupper L, Morgenstern H. *Epidemiologic Research. Principles and quantitative methods*. London: Lifetime Learning Publications, 1982.
4. Meinert C, Torascia S. *Clinical Trials: design, conduct and analysis*. Oxford: Oxford University Press, 1986.
5. Pocock S. *Clinical Trials: a practical approach*. Chichester: John Wiley & Sons, 1983.
6. Cook T., Campbell D. *Quasi-experimentation: design and analysis issues for field settings*. Boston: Houghton Mifflin company, 1979.
7. Morgenstern H. Uses of ecological analysis in epidemiologic research. *Am J Pub Health* 1982; 72:1336-1334.
8. Jenicek M, Cleroux R. *Epidemiología. Principios, técnicas, aplicaciones*. Barcelona: Salvat ed., 1987.
9. Breslow NE, Day NE. *Statistical methods in cancer research. Vol II: the design and analysis of cohort studies*. Lyon: IARC, 1987.
10. Schlesselman J. *Case-control studies: design, conduct, analysis*. Oxford: Oxford University Press, 1982.
11. Breslow NE, Day NE. *Statistical methods in cancer research. Vol I: the analysis of case-control studies*. Lyon: IARC, 1980.
12. Cornfield J. A method of estimating comparative rates from clinical data. *J Nat Cancer Inst* 1951; 11:1269-1275.
13. Ahlbom A. Study design and planning. In: Olsen J, Trichopoulos D (eds). *Teaching Epidemiology*. Oxford: Oxford University Press, 1992.
14. Rothman K. *Epidemiología Moderna*. Madrid: Díaz de Santos, 1987.
15. Sackett DL. Bias in analytic research. *J Chron Dis* 1979; 32:51-63.
16. Tomatis L (ed.). *Cancer: Causes, occurrence and Control* (IARC Sci Publ 100). Lyon: International Agency for Research on Cancer, 1990.
17. Weinberg C. Toward a clearer definition of confounding. *Am J Epidemiol* 1993; 137:1-8.

18. Alvarez-Dardet C, Bolúmar F, Porta M. La medición de la frecuencia de la enfermedad. *Med Clin* 1987; 88:287-291.
19. Jordan WM. Pulmonary embolism. *Lancet* 1961; 2:1146.
20. Centers for Disease Control. Pneumocystis carinii pneumonia - Los Angeles: *MMWR* 1981; 30:250.
21. Smith H, Spalding J Brote de parálisis en Marruecos debida al envenenamiento por ortofosfato de cresilo. *Lancet* II 1959; 1019-1021. En: Buck C, Llopis A, Nájera E, Terris M (eds). *El desafío de la Epidemiología*. Pub. Científica N°505. Washington, D.C.: OPS, 1988.
22. Comité Científico para el Síndrome del Aceite Tóxico. *El Síndrome del Aceite Tóxico. Conocimientos actuales y perspectivas*. Copenhague: Oficina Regional para Europa de la Organización Mundial de la Salud, 1992.
23. Alvarez-Dardet C, Mur P, Gascón E, Nolasco A, Bolumar F. La investigación clínica en España: tipos de diseños utilizados. *Med Clin* 1987; 89:221-223.
24. Hennekens Ch, Buring J. *Epidemiology in Medicine*. Boston: Little, Brown and Co., 1987.
25. Ministerio de Justicia e Interior. *Real Decreto 1332/1994*, de 20 de junio, por el que se desarrolla determinados aspectos de la Ley Orgánica 5/1992, de 29 de octubre, de regulación del tratamiento automatizado de los datos de carácter personal. Madrid: BOE 21 de junio de 1994.
26. Gili M, Vilches A, Garcia C, Cortes M, Rosado M, Nájera E. Bases epidemiológicas de los estudios ecológicos. *Rev San Hig Pub* 1986; 60: 237-250.
27. Cayuela A, Rodríguez S. *Atlas de mortalidad por cáncer. Andalucía 1980-1984*. Sevilla: Junta de Andalucía, 1993.
28. Cayuela A, Lacalle JR, Perea-Milla E, Martínez R, Velasco A, Gili M. Análisis de la mortalidad por cáncer de próstata. España 1951-1979. *Oncología* 1987; 10:364-370.
29. Cayuela A, Lacalle JR, Nieto C, Azpiri A, Gili M. Cardiopatía isquémica, neoplasias y factores de riesgo: correlación ecológica. *Rev Esp Cardiol* 1989; 42:84-89.
30. Cramer D. The role of cervical cytology in the declining morbidity and mortality of cervical cancer. *Cancer* 1974; 34:2018.
31. Briones E, Perea-Milla E, Ruiz-Canela MP, Torró C, Gili M. The Andalusian Nutritional Survey: comparison of the nutritional status of Andalusian children aged 6-60 months that of the NCHS/CDC reference population. *WHO Bull* 1989; 67:409-416.
32. Fernández R, Suárez C, Perea-Milla E, Vilaseca M, Díaz-Fierros M. Estudio de la calidad de la prescripción en la provincia de Málaga. *Gaceta Sanitaria* 1993 (suplemento); 7:17.
33. Infante-Rivard C, Fernández A. *Investigación epidemiológica en Salud Laboral*. Quebec: McGill University, 1994.
34. Poole Ch. Ecologic analysis as outlook and method (editorial). *Am J Public Health* 1994; 84:715-716.

35. Martínez C, Serrano S, Ortega R, Miñarro R, López B, Buendía A, Perea-Milla E. *Supervivencia del melanoma cutáneo en Granada*. XXII Congreso Nacional de Dermatología y Venereología. Granada: 1993.
36. Farmer AM. *Lecture notes on Epidemiology and Community Medicine*. Boston: Epidemiologic Resources Inc., 1983.
37. Doll R, Bradford Hill A. La mortalidad en relación con el hábito de fumar: diez años de observaciones sobre médicos británicos. *Br Med J* 1964; 1:1399-1410 y 1460-1467. En: Buck C, Llopis A, Nájera E, Terris M (eds). *El desafío de la Epidemiología*. Pub. Científica N° 505. Washington: OPS, 1988.
38. Riboli E. Nutrición y cáncer: antecedentes y base racional del European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann of Oncology* 1993; 2:121-129.
39. Arrighi HM, Metter EJ, Guess HA et al. Natural history of benign prostatic hyperplasia and risk of prostatectomy. *Urology* 1991; 38(suppl):4-8.
40. Thomas et al. 1983. Citado en: Sánchez-Cantalejo E. *Análisis de datos categóricos*. Granada: Escuela Andaluza de Salud Pública, 1990.
41. Nieto J, Peruga A. Riesgo Atribuible: sus formas, usos e interpretación. *Gaceta Sanitaria* 1990; 18:112-117.
42. Conzález CA, López-Abente G, Errezola M, Escolar A, Izarzugaza I, Riboli E, Nebot M. Diseño y realización de un estudio multicéntrico caso-control sobre cáncer de vejiga en España. *Med Clin* 1989; 92: 646-651.
43. Rogentine G, Yankee R, Gart J, Nam J, Trapati R. HLA antigens and disease: acute lymphocytic leukemia. *J Clin Invest* 1972; 51:2420-2428.
44. Rogentine G, Trapani R, Yankee R, Henderson E. HLA antigens and acute lymphocytic leukemia: the nature of the association. *Tissue Antigens* 1973; 3:470-476.
45. Norell SE. *A short course on Epidemiology*. New York: Raven Press Ltd., 1992.
46. Miettinen OS. *Theoretical Epidemiology. Principles of occurrence research in Medicine*. Albany NY: Delmar Publishers Inc., 1985.
47. Guallar EM. Alternativas al diseño tradicional de los estudios de cohortes: Estudios de casos y controles híbridos y estudios de caso-cohorte. *Revisiones en Salud Pública* 1991; 2:151-165.
48. Flanders W, Coleen A, Boyle C, Boring J. Bias associated with differential hospitalization rates in incident case-control studies. *J Clin Epidemiol* 1989; 42:395-401.
49. Gili M, Martínez-Cañabate E, Nieto C. Los estudios de casos y testigos. Los estudios transversales. *Enf Inf y Microbiol Clin* 1989; 7:559-563.
50. Martín JM. Oportunidad relativa: reflexiones en torno a la traducción del término 'odds ratio' (carta al director). *Gaceta Sanitaria* 1990; 4:37.
51. Pearce N. What does the Odds Ratio estimate in a case-control study?. *Int J Epidemiol* 1993; 22:1189-1192.
52. Walker AM. *Observation and inference. An introduction to the methods of Epidemiology*. Chestnut Hill MA: Epidemiologic Resources Inc., 1991.

53. Friedman G, Hiatt R, Quesenberry Ch, Selby J. Case-control study of screening for prostatic cancer by digital rectal examinations. *Lancet* 1991; 337:1526-1529.
54. Palli D, Roselli M, Buiatti E, Carli S, Ciatto S, Toscani L, Maltoni G. A case-control study of the efficacy of a non-randomized breast cancer screening program in Florence (Italy). *Int J Cancer* 1986; 38:501-504.
55. Davey G, Ströbele S, Egger M. Smoking and health promotion in Nazi Germany. *J Epidemiol Community Health* 1994; 48:220-223.
56. Gervás J, Pérez M.: El diseño semiexperimental y experimental y la investigación evaluativa. *Atención Primaria* 1991; 8:567-576.
57. Sánchez-Cantalejo Ramirez E. *Evaluación de intervenciones*. Granada: Escuela Andaluza de Salud Pública, 1991.
58. Veney J, Kaluzny D. *Evaluation and decision making for health services programs*. Englewoods Cliffs, NJ: Prentice Hall Inc., 1984.
59. Ministerio de Sanidad y Consumo. *Real Decreto 561/1993*, de 16 de abril, por el que se establecen los requisitos para la realización de ensayos clínicos con medicamentos. Madrid: BOE 13 de mayo de 1993.
60. *Directiva 91/507/CEE*, de 19 de julio. Unión Europea.
61. Legha SS. Ensayos clínicos en fase I en tumores sólidos en adultos. *Rev Cáncer* 1990; 4:11-14.
62. Martin M, Diaz-Rubio E. Estudios en fase II en Oncología. *Rev Cáncer* 1990; 4:15-23.
63. Lizón J. El estudio en fase III. Estudios comparativos aleatorios. *Rev Cáncer* 1990; 4:24-29.
64. Byar D, Piantados S. Factorial designs for randomized clinical trials. *Cancer Treatment Reports* 1985; 69:1055-1063.
65. Higgins J. *Introducción a los estudios clínicos de asignación aleatoria*. Durham, NC: Family Health International, 1985.
66. Skrabanek P, McCormick J. *Sofismas y desatinos en Medicina*. Barcelona: Ed. Doyma, 1992.
67. Fletcher R, Fletcher S, Wagner E. *Epidemiología Clínica*. Barcelona: Ed. Consulta, 1989.
68. Berry D. Interim analyses in clinical research. *Cancer Investigation* 1987; 5:469-477.

PARTE II
ANÁLISIS DE DATOS EN INVESTIGACIÓN
CLÍNICA Y EPIDEMIOLÓGICA

Emilio Sánchez Cantalejo
EDITOR

A. ANALISIS DESCRIPTIVO

E. Sanchez-Cantalejo Ramírez

R. Ocaña Riola

1. Introducción

Según L. E. Moses ¹, la estadística se puede definir como un cuerpo de conocimientos para aprender de la experiencia, frecuentemente en forma de números provenientes de medidas que muestran variaciones entre los distintos individuos. Cuando el clínico se fija en alguna característica de sus pacientes como pueda ser el sexo, la edad, el nivel de PSA (antígeno prostático específico), la presencia o no de adenoma de próstata, estadio del tumor, etc., lo primero que observa es que estas características no son iguales en todos sus pacientes; así no todos los pacientes tienen el mismo sexo, ni la misma edad, etc. Es decir, las características que se observan en la clínica cambian de unos pacientes a otros; a estas características se les denomina, por razones obvias, **variables**. Como, según se ha dicho antes, la estadística es la ciencia que estudia la variabilidad, no es sorprendente que ésta sea una herramienta útil para el clínico para avanzar en sus conocimientos; por otra parte, cualquier lector de publicaciones científicas se ve abocado, si quiere mantenerse al día en su materia, a tener unos conceptos estadísticos mínimos que le hagan comprender algunas cosas «raras» que cada vez aparecen con más frecuencia en la literatura médica; a saber: media, error estandar, O.R., histograma, valor P, análisis multivariante, etc. De la misma manera que el clínico necesita cada día utilizar una tecnología más sofisticada tanto para diagnosticar como para tratar las enfermedades de sus pacientes y ello le supone un esfuerzo de puesta al día, también se ve en la necesidad de tener una cultura estadística que le permita al menos poder leer lo que publican las revistas más importantes de su especialidad.

Es impresionante el aumento del uso de los procedimientos estadísticos en las revistas clínicas más prestigiosas; por ejemplo, en una revisión de los 115 artículos originales publicados en el volumen 321 (1989) de la revista *The New England Journal of Medicine*, el 88% de ellos utilizaba algún método estadístico distinto a los puramente descriptivos y, por otra parte, el número medio de métodos estadísticos por artículo es casi tres ². Es cada vez más frecuente la presencia de revisores estadísticos en los comités editoriales de revistas clínicas; el *British Medical Journal* es una de

ellas y ha publicado la última versión³ de las instrucciones para los autores que deseen publicar en tal revista. En tal referencia se dice que todo artículo a publicar pasa por uno de los siete estadísticos pertenecientes al comité editorial y también aparece el listado de preguntas acerca del diseño y análisis del estudio candidato a publicación. En nuestro país, aunque con algún retraso, revistas como *Medicina Clínica* y *Gaceta Sanitaria* han incorporado a estadísticos en el proceso de revisión de artículos. Este aumento del uso de los métodos estadísticos ha llevado aparejado en muchas ocasiones un mal uso de los mismos⁴; algún autor ha cifrado en el 50% la proporción de artículos clínicos en los que se encuentran serios problemas de diseño y/o análisis de los datos. Altman⁵ escribe un editorial alertando sobre este problema donde establece un parangón entre el clínico que no utiliza el tratamiento correcto en sus pacientes con el investigador clínico que no diseña y/o analiza de forma adecuada sus estudios. Las razones de esta situación son, por una parte, el poco conocimiento por parte de los clínicos de disciplinas como la estadística y la epidemiología y, por otra, la falta de profesionales de estas disciplinas, problema especialmente grave en nuestro país.

Son multitud el número de publicaciones periódicas biomédicas, tanto generales como especializadas, que han publicado artículos sobre metodología tanto estadística como epidemiológica, con el objetivo de aumentar la cultura científica entre sus lectores. Esta parte de la monografía trata de introducir al clínico en los conceptos básicos de análisis de datos más comunmente utilizados en la literatura biomédica.

2. Medición de signos y síntomas en clínica

El clínico cuando examina a un paciente recoge gran cantidad de información acerca de sus características con objeto de mejorar tanto el diagnóstico como el tratamiento de la enfermedad; como quiera que tales características cambian de unos pacientes a otros, de ahora en adelante las vamos a denominar variables. Está claro que no todas esas variables son de igual naturaleza; podemos medir variables como la presencia de dolor en hipocondrio derecho, o bien podemos medir el nivel de PSA; los posibles valores de la primera variable los podemos nombrar por «si» y «no», por «presencia» y «ausencia», etc; sin embargo, los posibles valores de la segunda variable son cualquier número entre, por ejemplo, 0 y 2000 ng./ml.; podemos decir que los posibles valores de la primera son etiquetas, mientras los de la segunda son números. Clásicamente las variables se han clasificado en dos grandes grupos: las **variables cualitativas** y las **cuantitativas**.

Se dice que una variable es cualitativa o categórica si sus posibles valores no son numéricos; ejemplos de este tipo de variables son: el sexo, el estado civil, el estadio del tumor, la variable GLEASON (gradación de la malignidad de un tumor), etc.; todas

ellas tienen en común el hecho de que sus valores no son números. Del sexo se dice que es una **variable dicotómica o binaria** pues solo puede tomar dos valores: hombre, mujer; sin embargo, el estado civil es también categórica pero sus posibles valores son más de dos: casado, soltero, divorciado, etc.; a estas variables se les denomina **policotómicas**. En la variable estadio de cáncer y en la GLEASON hay algo que no se encuentra en las otras y esto es que entre sus valores o categorías hay implícito un orden: un tumor en estadio I precede y tiene mejor pronóstico que en estadio II, éste mejor que el III, etc.; a estas variables se les denomina **ordinales**; sin embargo, aún no se ha «demostrado», a pesar de algunos, que entre los valores de la variable sexo exista un orden. El hecho de que las variables categóricas no sean numéricas no entra en contradicción con el hecho de que para poder manejarlas con muchos paquetes informáticos sea necesario asignar números a cada valor; así podemos asignar valor 1 al paciente casado, 2 al soltero, 3 al separado, etc., aunque a la hora de analizar los datos deberemos advertir a la máquina que tales números son realmente etiquetas.

El grupo de las cuantitativas está constituido por aquellas variables cuyos valores son numéricos: número de hijos, edad, nivel de PSA, etc. La primera se dice que es una variable cuantitativa **discreta** pues sus posibles valores son números aislados: 0, 1, 2, etc.; tanto la edad como el nivel de PSA son variables cuantitativas **continuas** porque, al menos teóricamente, pueden valer cualquier número, por raro que sea, dentro de un intervalo determinado.

El distinguir entre los distintos tipos de variables no es gratuito, pues el análisis a efectuar con las variables dependerá de su naturaleza; así, dependiendo de si la variable es categórica, cuantitativa discreta o continua, utilizaremos distintas estrategias de presentación y análisis.

3. Presentación de datos

Una vez medidas las variables de interés en los sujetos del estudio, es necesario resumir de alguna manera la información recogida. La estadística descriptiva es un conjunto de procedimientos encaminados a sintetizar tal información; entre éstos podemos citar: las tablas, gráficos y resúmenes numéricos acerca de la magnitud y variabilidad de las mediciones realizadas.

3.1. Tablas

Si pretendiéramos realizar un análisis estadístico sobre la gravedad de una serie de pacientes afectados de cáncer de próstata, es claro que un listado del grado de

malignidad del tumor para cada una de estas personas no nos daría idea de la distribución del estado de los individuos; de ahí la conveniencia de resumir adecuadamente esa información. Una forma de presentar resumidamente esos datos es mediante una **tabla**. Una tabla se construye poniendo en una columna las categorías o valores de la variable en cuestión, en este caso el grado de malignidad del tumor dado por la variable GLEASON, y en otra columna el número de veces que se repite cada una de esas categorías; en la Tabla II.1 aparece la distribución de esta variable medida en 74 pacientes con cáncer de próstata.

Tabla II.1.- Distribución del grado de malignidad del tumor en 74 individuos con cáncer de próstata.

GLEASON	Frecuencia absoluta	Frecuencia relativa
Grado 4	12	0.1622
Grado 5	18	0.2432
Grado 6	21	0.2838
Grado 7	14	0.1892
Grado 8 o más	9	0.1216
Total	74	1.0000

Según esta tabla, de entre las 74 personas en estudio, 12 presentan un tumor considerado de Grado 4; este número 12 se dice que es la **frecuencia absoluta**, o simplemente la **frecuencia**, de la categoría «Grado 4», pues es el número de veces que se repite esa categoría; por otra parte, el 12.16% de los pacientes están clasificados como individuos con un tumor de Grado 8 o superior, es decir, la **frecuencia relativa** de la categoría «Grado 8 o más» es 0.1216. En general, la frecuencia de una categoría es el número de veces que se presenta tal categoría, mientras que la frecuencia relativa es el cociente entre la frecuencia absoluta y el número total de individuos; es claro que la frecuencia relativa depende del conjunto de referencia, en este caso las personas diagnosticadas con cáncer de próstata en una determinada región, mientras que la frecuencia absoluta es independiente de éste; si notamos por f_i la frecuencia absoluta de la clase i , la frecuencia relativa, que notaremos por h_i , se puede expresar mediante el cociente

$$h_i = \frac{f_i}{n}$$

siendo n el número de individuos del conjunto de referencia.

Las distintas categorías en que se puede clasificar a los individuos, y esto vale para cualquier tipo de variable, se deben construir según dos criterios: **exhaustividad** y **exclusividad**; en otras palabras, un individuo debe pertenecer a **una y sólo a una** de las categorías. Estos criterios implican las siguientes relaciones:

$$f_1 + f_2 + \dots + f_k = n$$

$$h_1 + h_2 + \dots + h_k = 1$$

que resumidamente se escribe así

$$\sum_{i=1}^k f_i = n \qquad \sum_{i=1}^k h_i = 1$$

donde el símbolo $\sum_{i=1}^k$ significa que hay que sumar desde la categoría 1 hasta la categoría k , siendo k el número de categorías, que para la tabla anterior vale 5.

Para variables cuantitativas, el proceso de construcción de la tabla es parecido; para el caso discreto, la tabla se compone de una columna donde aparecen los distintos valores de la variable y en otra sus frecuencias, como se muestra en la Tabla II.2 de la distribución de la variable «Número de sesiones de tratamiento necesarias para la fragmentación total» en un conjunto de 200 pacientes tratados mediante litotricia extracorpórea.

Tabla II.2.- Distribución del número de sesiones de tratamiento en 200 pacientes sometidos a tratamiento mediante litotricia.

Nº de sesiones	Frecuencia absoluta	Frecuencia relativa
1	165	0.825
2	26	0.130
3	7	0.035
4 o más	2	0.010
Total	200	1.000

Para el caso de una variable continua como, al menos teóricamente, puede tomar un número muy grande de valores, en una columna no se representan cada uno de los posibles valores, sino **intervalos de valores**; un intervalo viene dado por una pareja

de valores que son los **extremos** del intervalo; a la distancia entre los extremos, su diferencia, se denomina **amplitud** del intervalo; a los efectos que después consideraremos, es necesario definir lo que llamaremos la **marca de clase** del intervalo, como el punto intermedio entre sus dos extremos, y que en ocasiones tomaremos como representante de todos los valores del intervalo correspondiente. En la Tabla II.3 se muestra la distribución de la duración del tratamiento de cada litiasis, dada en minutos, para el conjunto de los 200 pacientes mencionados anteriormente. La variable considerada es continua.

Tabla II.3.- Distribución de la duración del tratamiento en 200 pacientes sometidos a litotricia.

Duración (en minutos)	Frecuencia absoluta	Marca de clase	Frecuencia acumulativa absoluta	Frecuencia Acumulativa relativa
$10 \leq x < 20$	4	15	4	0.0200
$20 \leq x < 30$	5	25	9	0.0450
$30 \leq x < 40$	21	35	30	0.1500
$40 \leq x < 50$	27	45	57	0.2850
$50 \leq x < 60$	48	55	105	0.5250
$60 \leq x < 70$	53	65	158	0.7900
$70 \leq x < 80$	28	75	186	0.9300
$80 \leq x < 90$	13	85	199	0.9950
$90 \leq x < 100$	1	95	200	1.0000
Total	200	1.0000		

En caso de categorías ordenadas, es decir, cuando se trate de variables categóricas ordinales o variables cuantitativas, tiene sentido definir lo que se llama **frecuencia acumulativa absoluta y/o relativa** de la categoría como la frecuencia absoluta y/o relativa propia más las de todas las categorías anteriores; las frecuencias acumulativas para la distribución de la duración del tratamiento de cada litiasis aparecen en las dos últimas columnas de la tabla; que la frecuencia acumulativa del intervalo $50 \leq x < 60$ sea 105 significa que de entre los 200 pacientes, 105 fueron tratados durante un periodo de tiempo inferior a 60 minutos; que su frecuencia acumulativa relativa sea 0.525 significa que el 52.5% de los pacientes tuvieron un tratamiento cuya duración fue inferior a 60 minutos.

3.2. Gráficos

La representación gráfica de las observaciones realizadas es una forma sencilla y rápida de comunicar información. Los gráficos son una herramienta estadística con

mucha solera; Playfair, uno de los fundadores de los gráficos estadísticos, propuso a principios del siglo XIX el uso de los diagramas de barras y el de sectores para representar sus datos. Otra gran precursora de la utilización de la estadística en el campo sanitario, Florence Nightingale, también utilizaba gráficos para presentar las bajas mensuales del ejército británico en la guerra de Crimea.

Los gráficos son un medio extraordinariamente importante para comunicar información en artículos científicos; análogamente al lenguaje escrito que puede comunicar ideas, teorías, números, etc., el lenguaje gráfico es muy utilizado para mostrar información porque es un medio muy efectivo. Tendencias y relaciones en nuestras observaciones se evidencian por medio de gráficos debida a la enorme potencia de nuestro sistema ojo-cerebro para percibirlos; este sistema puede resumir gran cantidad de información, percibir sus características más sobresalientes o fijarse en detalles que no se pueden captar con la información numérica. La reciente explosión tanto de *hardware* como de *software* nos permite ahora poder hacer un uso más fácil de las representaciones gráficas y es la razón por la que constantemente se están haciendo nuevas propuestas de representaciones gráficas que mejoren o complementen las ya existentes. Revistas como el *Journal of Clinical Investigation* y *The Lancet* dedicaban ya en 1981 un 12% y un 10% de la superficie de sus artículos originales a representaciones gráficas; el lector interesado en estos temas puede consultar la bonita obra de Tufte⁶.

Como todo método que se usa con frecuencia, no siempre se utiliza de la mejor forma, por lo que existen unas normas que se deben tener en cuenta a la hora de construir toda representación gráfica:

- 1) Los gráficos deben ser visualmente claros.
- 2) Deben estar claramente descritos tanto en el pie de la figura como en el texto del artículo.
- 3) Siempre que sea posible, se deben representar gráficamente las conclusiones más importantes que se derivan del estudio.
- 4) Evitar los gráficos confusos; por ejemplo, no sobrecargar mucho la gráfica de tal manera que «los árboles no dejen ver el bosque».

El método gráfico a utilizar depende tanto del número de variables a representar como de la naturaleza de éstas. De todas formas, el principio general subyacente en la mayoría de los métodos gráficos consiste en asignar a cada valor o grupo de valores de las variables áreas proporcionales a su frecuencia; a continuación se presentan los procedimientos más usuales.

3.3. Gráficos para datos unidimensionales

Son los gráficos a utilizar cuando se quiere representar una sola variable; quizás el más utilizado en la literatura médica sea el llamado **diagrama de barras**. Tanto las variables cualitativas como las cuantitativas discretas se pueden representar mediante este tipo de gráfico, que se construye por medio de un sistema cartesiano en el que en el eje horizontal se representan los valores de la variable y por cada uno de esos valores se trazan segmentos o barras perpendiculares cuya altura sea proporcional a la frecuencia, absoluta o relativa, de cada valor. La Fig. II.1 muestra la distribución del grado de malignidad, medido en términos de GLEASON, en 74 pacientes con cáncer de próstata que aparecen en la Tabla II.1; como quiera que, aunque se le asignen números a las distintas categorías del grado de malignidad, tal variable no es numérica, pues si se acepta su carácter numérico equivaldría a afirmar que existe la misma diferencia entre un grado 4 y un 5 que entre éste y un 6 etc.; a cada categoría de la variable se le ha asignado un rectángulo con igual base de valor arbitrario, pero igual para todas las categorías; sobre cada una de esas bases se levantan rectángulos con altura proporcional a su frecuencia, con lo que se consigue que sus áreas sean proporcionales a sus frecuencias.

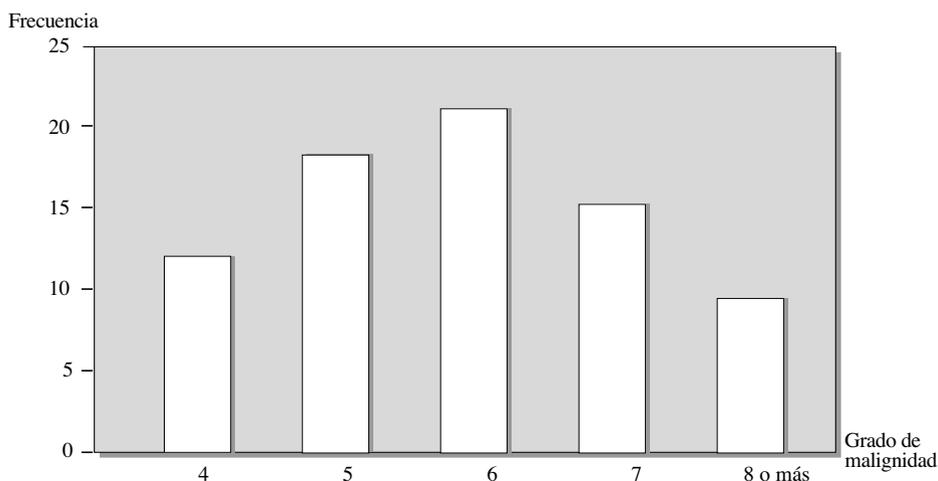


FIG. II.1 Distribución del grado de malignidad del tumor en 74 casos de cáncer de próstata.

Cuando se quiere representar la distribución de una variable en más de una serie de datos, es decir, cuando se quieren comparar dos o más distribuciones, se pueden hacer gráficas conjuntas para las distintas series; en la Fig. II.2 aparecen las distribuciones del grado de malignidad del cáncer de próstata en 17 individuos con afectación tumoral de los ganglios linfáticos y 57 sin afectación, en la que se puede apreciar

cómo la mayoría de los casos de cáncer de próstata sin afectación ganglionar tienen grados de malignidad menores que los casos con afectación ganglionar.

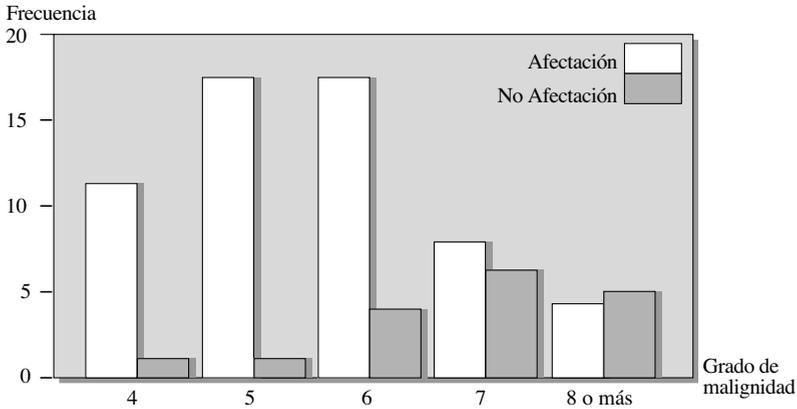


FIG. II.2.- Distribución del grado de malignidad del tumor según la afectación ganglionar.

Para el caso de variables continuas, en el eje horizontal se toman segmentos que sean proporcionales a las amplitudes de los intervalos y, con esas bases, se trazan rectángulos tales que su área sea proporcional a la frecuencia del correspondiente intervalo; una representación de este tipo se denomina **histograma**. Uniendo los puntos medios de las bases superiores de estos rectángulos se consigue el **polígono de frecuencias**. En la Fig. II.3 aparece la distribución de la duración del tratamiento con litotricia en un grupo de 200 pacientes, expresada en minutos.

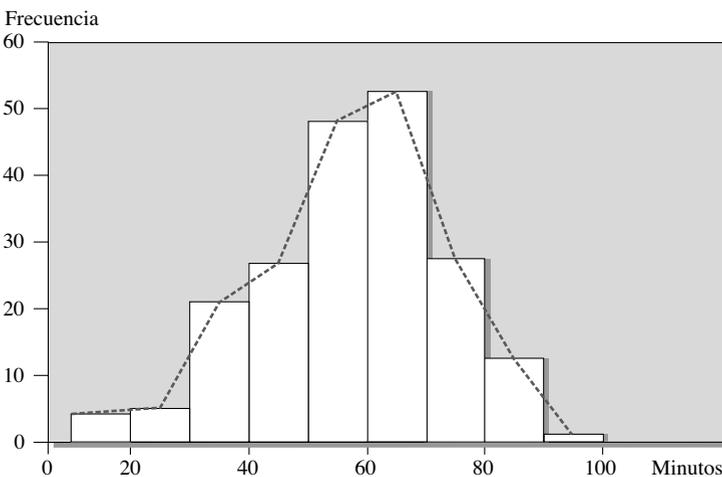


FIG.II.3.- Distribución de la duración del tratamiento mediante litotricia extracorpórea en 200 pacientes.

Ya que el área del rectángulo es el producto de su base por su altura, eligiendo intervalos de igual amplitud, el área dependerá sólo de la altura lo que facilitará la comprensión de la representación; de ahí que, siempre que sea posible, se deben elegir intervalos de valores de igual amplitud.

Con las frecuencias acumulativas se puede construir los histogramas de frecuencias acumulativas, tanto absolutas como relativas; a modo de ejemplo, en la Fig. II.4 aparece el **polígono de frecuencias acumulativas absolutas** para la distribución de la duración del tratamiento de litiasis en 200 pacientes; este polígono nos permite evaluar gráficamente la frecuencia acumulativa, en este caso absoluta, para cualquier valor de la variable representada, en este caso la duración del tratamiento. Más adelante se hará uso del polígono de frecuencias acumulativas relativas.

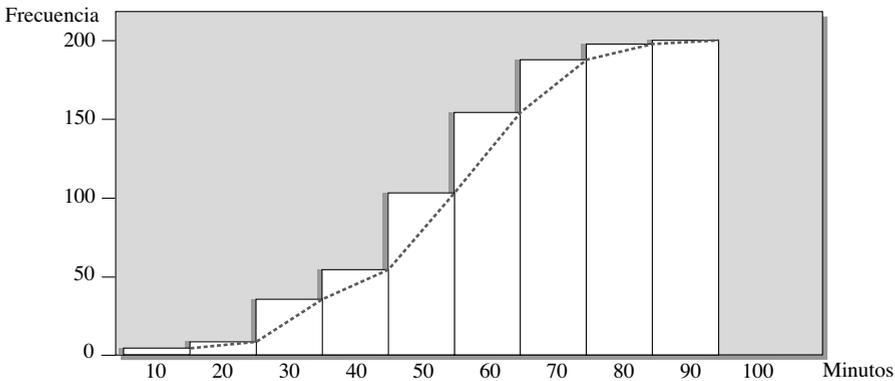


Fig.II.4.- Histograma y polígono de frecuencias acumulativas absolutas de la duración del tratamiento mediante litotricia extracorpórea.

Aparte de éstas, también se utilizan con frecuencia otros tipos de representaciones gráficas; entre las más comunes están los **diagramas de sectores** o **tartas** que se construyen asignando a cada valor de la variable un sector circular cuya área es proporcional a la frecuencia de tal valor, como aparece en la Fig. II.5.

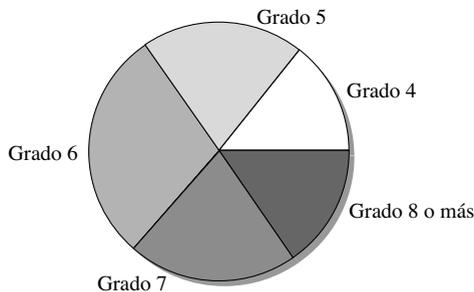


Fig.II.5.- Distribución del grado de malignidad del tumor

Una representación gráfica menos conocida que las anteriores es el llamado «**tronco y hoja**», traducción de su nombre original «stem-and-leaf»; esta representación, debida a J. W. Tukey, que podemos considerar como un híbrido entre una tabla y un histograma, tiene como ventaja sobre estas dos el hecho de, aparte de representar la forma de la distribución, también muestra los valores de la variable. La idea consiste en dividir cada dato de la serie en dos partes: una que es «el tronco» y otra «la hoja». A continuación aparecen las presiones sistólicas de un grupo de 80 adultos.

105	110	141	120	112	119	117	123
142	138	99	121	113	128	112	128
109	133	133	126	112	125	113	127
105	120	108	118	117	127	119	120
122	114	115	122	110	100	123	129
122	123	108	112	106	111	101	128
117	111	107	113	104	118	104	121
118	120	117	135	106	119	123	122
119	107	121	139	121	122	127	128
129	108	120	119	118	128	117	106

La siguiente representación es el tronco y hoja de los datos anteriores

```

09 | 9
10 | 5 9 5 7 8 8 8 7 6 4 6 0 1 4 6
11 | 7 8 9 0 4 1 5 7 8 2 3 9 2 3 2 7 0 8 9 1 8 9 7 2 3 9 7
12 | 2 2 9 0 3 0 1 0 0 1 6 2 1 8 5 7 2 8 3 3 7 3 8 7 0 9 8 1 2 8
13 | 8 3 3 5 9
14 | 2 1

```

Las cifras a la izquierda de la vertical son los troncos y las de la derecha, las hojas; el primer valor 105 está representado con tronco 10 y hoja 5; el valor 142 como 14 y 2, etc. Por tanto, el tronco 10 tendrá tantas hojas como presiones haya comprendidas entre 100 y 109 mm; es decir, la longitud de cada fila de hojas es el número de veces que se repite cada tronco. Como se ve, con esta representación gráfica obtenemos información acerca de la forma de la distribución, la que nos daría el histograma, pero además se tiene la información que nos daría una tabla pues permite conocer los valores de la variable, cosa que no ocurriría con el histograma. De la misma manera que en el histograma se puede elegir la amplitud de los intervalos que sirven de base a los rectángulos, para esta representación gráfica es también posible adecuar tal amplitud a nuestra conveniencia. El anterior tronco y hoja se ha construido en base a intervalos de amplitud 10 mm.; si por alguna razón nos interesase afinar un poco más y construirlo en base a intervalos de 5 mm., la representación gráfica sería así

09 | 9
 10 | 4 0 1 4
 10 | 5 9 5 7 8 8 8 7 6 6 6
 11 | 0 4 1 2 3 2 3 2 0 1 2 3
 11 | 1 7 8 9 4 5 7 8 9 7 8 9 8 9 7 9 7
 12 | 2 2 0 3 0 1 0 0 1 2 1 2 3 3 3 0 1 2
 12 | 1 9 6 8 5 7 8 7 8 7 9 8 8
 13 | 3 3
 13 | 8 5 9
 14 | 2 1

en donde en el primer tronco 10 estarían las presiones comprendidas entre 100 y 104 mm., ambas inclusive, y en el otro tronco 10 se situarán las presiones entre 105 y 109 mm.

3.4. Gráficos para datos bidimensionales

Hasta ahora se han hecho representaciones gráficas de una variable medida en varios individuos; sin embargo, muchas veces se miden más de una variable en cada sujeto y suele ser de interés estudiar la relación entre dos de ellas; la representación gráfica de las parejas de valores es un paso previo y muy importante a la hora de estudiar relaciones entre dos variables. La situación general es la de tomar dos medidas en cada individuo y tratar de conocer como varía una de las variables cuando varía la otra. Ya que tenemos el valor de dos variables para cada individuo y como una pareja de valores se puede identificar como un punto en el plano, en definitiva la representación gráfica vendrá dada por n puntos en el plano. En la siguiente tabla aparecen los trasplantes renales efectuados en Andalucía entre los años 1986 a 1993.

Tabla II.4.- Trasplantes de riñón en Andalucía 1986-1993.

Año	1986	1987	1988	1989	1990	1991	1992	1993
Trasplantes	183	144	144	124	153	161	206	187

Aquí el interés está en representar gráficamente como ha evolucionado el número de trasplantes de riñón durante el periodo 1986-1993 en Andalucía. Veamos como hacerlo; para ello en un sistema de dos ejes representamos en el eje de abscisas, el horizontal, el tiempo y en el eje de ordenadas, el vertical, el número de trasplantes realizados, como aparece en la Fig. II.6 (a). Este tipo de representación gráfica, denominada

nube de puntos, es fácil de manipular; en la Fig. II.6 (b) aparece la misma representación en donde se ha cambiado la escala del eje vertical con lo que da la sensación de menos cambios interanuales en el número de trasplantes renales en Andalucía. Las dos gráficas son perfectamente válidas pero evidentemente la impresión que causan al lector es muy diferente. De este hecho se puede sacar la conclusión de que es muy importante, a la hora de mirar cualquier nube de puntos, fijarse en las escalas elegidas, evitando de esa manera cualquier posible manipulación malintencionada.

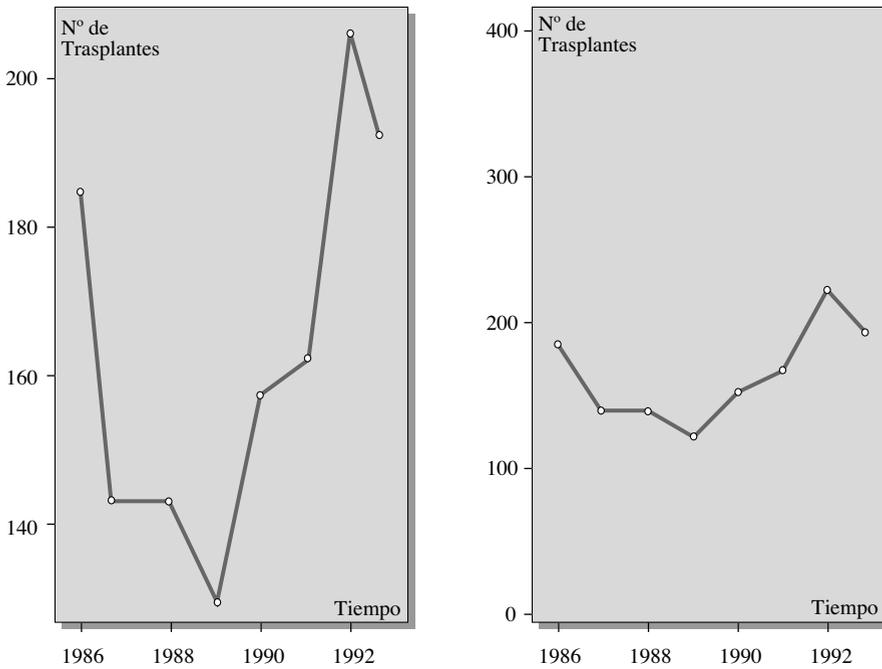


Fig. II.6.- Trasplantes en Andalucía en el periodo 1986-1993. La diferencia entre los gráficos (a) y (b) está en la escala del eje vertical.

También es posible estudiar la relación de dos o más variables con otra tercera. A continuación aparece el número de trasplantes de corazón en la misma época anterior.

Tabla II.5.- Trasplantes de corazón en Andalucía. 1986-1993

Año	1986	1987	1988	1989	1990	1991	1992	1993
Trasplantes	3	3	10	12	24	43	50	51

Una posible estrategia consiste en construir dos gráficos, uno para el número de trasplantes renales y otro para los cardiacos; sin embargo, puede ser más informativo un gráfico que presente simultáneamente ambas relaciones, como aparece en la Fig. II.7, en donde se observan las fluctuaciones en el número de trasplantes renales y el aumento en los cardiacos a lo largo del periodo considerado.

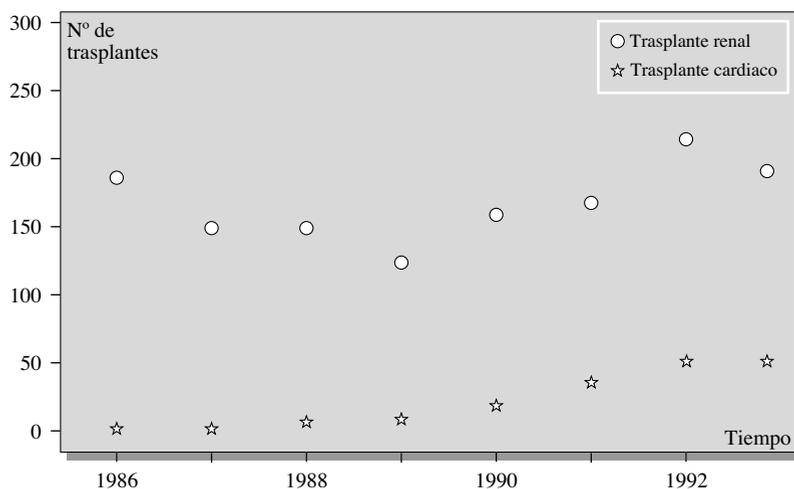


Fig. II.7.- Trasplantes renales y cardiacos en Andalucía en el periodo 1986-1993.

3.5. Gráficos para datos multidimensionales

Es en el caso en que para cada observación se han medido tres o más variables, donde ha habido novedades más importantes en cuanto a métodos gráficos de representación de datos; la mayoría de los métodos sólo se pueden llevar a cabo con las capacidades de los ordenadores, en especial, algunos paquetes gráficos tales como SAS, STATGRAPHICS o S-PLUS. Entre las distintas técnicas sólo vamos a citar las que quizás son más utilizadas: los **diagramas de estrellas** y las **caras de Chernoff**.

Un diagrama de estrellas consiste en asignar a cada sujeto una estrella con tantos radios como variables se hayan medido; es un tipo de gráfico pensado para variables cuantitativas siendo la longitud de cada radio de la estrella proporcional a la magnitud de la variable que representa. Como observaciones para construir un diagrama de estrellas vamos a tomar las que aparecen en la Tabla II.6, donde se muestra una serie de cinco indicadores que tienen que ver con la gestión, en seis grandes hospitales andaluces; por tanto, cada hospital vendrá representado por una estrella de cinco radios que tendrán longitud proporcional a los valores de las cinco variables.

Tabla II.6.- Indicadores en 6 grandes hospitales andaluces.

Hospital	Estancia media	Índice de ocupación	Utilización de quirófano	Lista de espera	Mortalidad
Hospital 1	10.60	87.25	57.88	128.15	35
Hospital 2	8.91	74.58	69.99	196.10	26
Hospital 3	11.01	89.27	66.33	65.15	31
Hospital 4	11.19	83.22	70.70	134.87	26
Hospital 5	8.95	85.46	66.77	113.50	28
Hospital 6	9.06	72.08	65.58	66.21	34

Como quiera que entre estas variables medidas hay algunas cuyos valores están en relación inversa y otras en relación directa a la calidad de la gestión del hospital, realizaremos alguna transformación. Por ejemplo, ya que a mayor valor de la variable «estancia media», peor gestión, utilizaremos la variable «12-estancia media» por lo que a mayor valor de esta variable, mejor gestión del hospital; de la misma forma utilizaremos «200-lista de espera» y «36-mortalidad». Con estas transformaciones, las cinco variables son homogéneas en el sentido de que cuanto mayores sean sus valores mejor gestión.

El programa utilizado para el diagrama de estrellas ha sido S-PLUS el cual asigna a la primera variable, en este caso «12-estancia media», el eje horizontal del primer cuadrante, asignando a las restantes variables los sucesivos ejes siguiendo el sentido contrario al de las agujas del reloj. Se ha de tener en cuenta que tiene sentido comparar longitudes de un mismo eje en las distintas estrellas; sin embargo, dentro de una misma estrella los distintos ejes miden diferentes escalas y no tiene sentido compararlos. De la Fig. II.8 se desprende que el Hospital 5 es, en conjunto, el mejor gestionado y que el Hospital 1 es el peor, en relación a las 5 variables medidas. Obsérvese como para cada una de las variables, al menor valor se le asigna un radio de longitud cero y a los restantes valores radios de longitud proporcional a su diferencia con el valor menor.

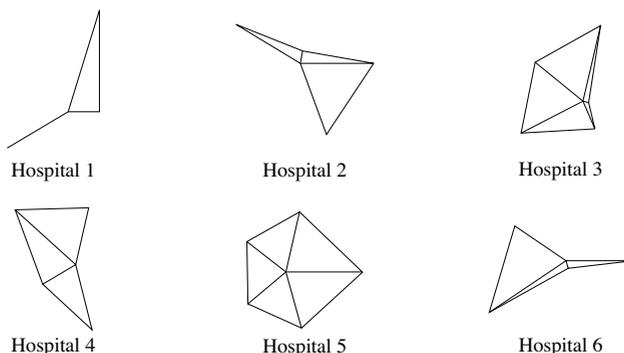


Fig. II. 8.- Diagrama de estrellas para cinco indicadores de gestión en seis hospitales andaluces.

Otro método de representación de observaciones multidimensionales es el propuesto por Chernoff en 1973 consistente en asignar a cada variable una característica de la cara y ya que el ojo humano es bastante experto en distinguir entre caras con distintas características, es un método muy útil de presentar información en muchas dimensiones. Cada programa informático permite un número de variables a representar; S-PLUS permite hasta quince variables siendo las cinco primeras las representadas por el área de la cara, forma de la cara, longitud de la nariz, localización de la boca y curvatura de la sonrisa. A continuación aparece este gráfico para los seis hospitales.

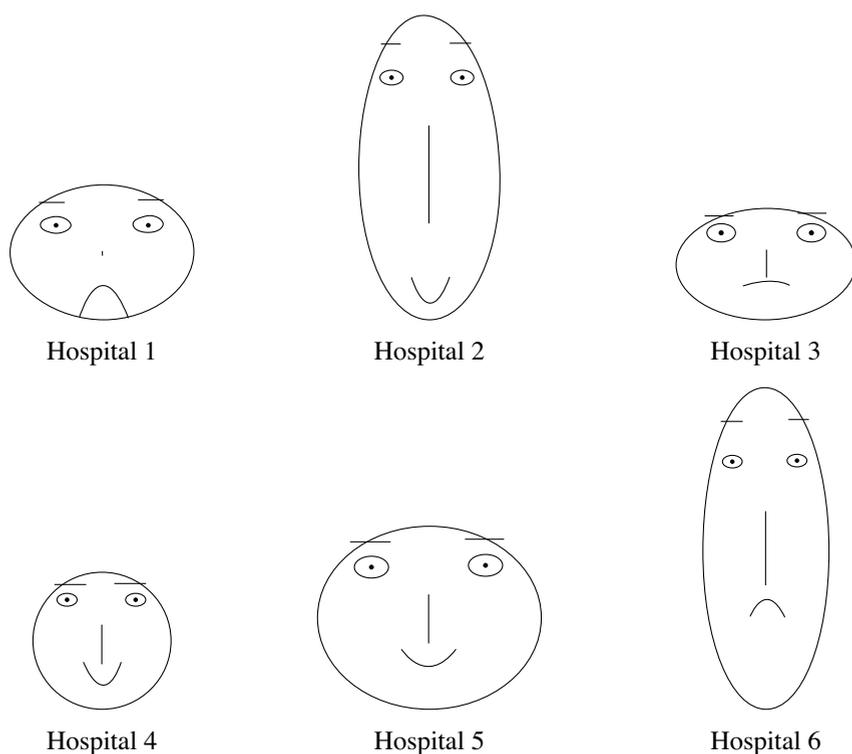


Fig. II.9.- Caras de Chernoff para cinco indicadores de gestión en seis hospitales andaluces.

Las variables utilizadas han sido introducidas de manera que valores pequeños de ellas representan una mala gestión hospitalaria, lo que se traduce en rasgos faciales tristes y pequeños; de esta forma es fácil apreciar que en general, los peor gestionados son los hospitales 1, 3 y 6 mientras que los mejor gestionados son los hospitales 2, 4 y 5. Para comprender el significado de las caras puede ser útil la siguiente tabla donde se explican las características de cada rasgo:

Rasgo facial	Característica	Significado
Área de la cara	<ul style="list-style-type: none"> • Grande • Pequeña 	<ul style="list-style-type: none"> - Estancia media pequeña - Estancia media elevada
Forma de la cara	<ul style="list-style-type: none"> • Achatada • Alargada 	<ul style="list-style-type: none"> - Índice de ocupación elevado - Índice de ocupación pequeño
Longitud de la nariz	<ul style="list-style-type: none"> • Grande • Pequeña 	<ul style="list-style-type: none"> - Gran uso de quirófano - Escaso uso de quirófano
Localización de la boca	<ul style="list-style-type: none"> • Cercana a la nariz • Alejada de la nariz 	<ul style="list-style-type: none"> - Lista de espera pequeña - Lista de espera elevada
Curvatura de la sonrisa	<ul style="list-style-type: none"> • Sonriente • Triste 	<ul style="list-style-type: none"> - Escasa mortalidad - Elevada mortalidad

Observemos que la cara correspondiente al hospital 5 es la más sonriente y proporcionada, por lo que puede decirse que, en general, es el hospital 5 el mejor gestionado, aunque existan características puntuales como la cercanía de la boca a la nariz, representativa del tamaño de la lista de espera, que ha sido mejorada por el hospital 3.

4. Resumen numérico de una serie estadística

Además de las tablas y los gráficos, una serie de observaciones se puede describir y resumir mediante lo que denominaremos **estadísticos**; un estadístico no es más que una forma de combinar las observaciones; dicho más formalmente, un estadístico es una función de los datos observados. Existe gran variedad de estadísticos que se pueden usar para describir una colección de datos y los hay de dos tipos:

- 1) Unos dan idea de la magnitud o tamaño de los datos; son las llamadas **medidas de posición o tendencia central**.
- 2) Otros nos proporcionan información acerca de la heterogeneidad de nuestras observaciones; son las conocidas como **medidas de dispersión o variabilidad**.

4.1. Medidas de tendencia central

La medida de tendencia central más usual es la **media aritmética** o, simplemente, la **media**; éste es un estadístico que se calcula como la suma de todas las observa-

ciones dividida por el número de ellas. En la Tabla II.7 aparece el coste por receta, en pesetas corrientes, en el segundo semestre de 1992, en Andalucía.

Tabla II.7. Coste medio por receta en Andalucía, Julio-Diciembre 1992.

Mes	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Coste/receta	1170	1174	1188	1212	1199	1189

Para calcular el coste medio por receta en ese periodo no hay más que sumar los costes y dividir la suma por 6, el número de meses, que es el número de observaciones realizadas.

$$\text{Coste medio por receta} = \frac{1170 + \dots + 1189}{6} = \frac{7132}{6} = 1188.67$$

donde 1188.67 es un número que resume los seis valores del coste por receta en el periodo indicado. Como toda medida de tendencia central, es un valor que está por el centro de la distribución, en el sentido de que no es ni de los valores bajos ni de los altos del conjunto.

Consideremos ahora los datos de la Tabla II.8, en donde se muestra el gasto farmacéutico por habitante en 1992 de los diferentes distritos sanitarios de la provincia de Granada.

Tabla II.8.- Prestaciones farmacéuticas en los distritos sanitarios de Granada.

Distrito	Población	Gasto total (en miles)	Gasto por habitante
Distrito 1	73423	944954	12870
Distrito 2	57110	702510	12301
Distrito 3	91771	1341967	14623
Distrito 4	40479	692393	17105
Distrito 5	202758	3128556	15430
Distrito 6	65052	972983	14957
Distrito 7	217772	3255474	14949
Distrito 8	74937	1229192	16403
Total	823302	12268029	118638

Para conocer la media del gasto farmacéutico por persona en el conjunto de los ocho distritos sanitarios, no habría más que sumar todos estos gastos a nivel de distrito y dividirlos por 8, el número de distritos.

$$\text{Coste medio/persona} = \frac{12870 + \dots + 16403}{8} = 14830$$

Sin embargo, la misma fuente asegura que el gasto total en recetas fué de 12268029000 pesetas y que la población protegida fueron 823302 personas; por tanto el gasto medio por persona también se podría calcular así:

$$\frac{12268029000}{823302} = 14901$$

¿A qué se debe esta discrepancia en los resultados?. El primer gasto medio por receta calculado es erróneo y lo es porque, evidentemente, no todos los distritos tienen la misma importancia, entendida ésta como el número de personas protegidas. Es claro que al gasto medio de distritos como el 5 o el 7 habría que darle más peso, más importancia, que al gasto medio de distritos como el 2 o el 4 que tienen muchas menos personas protegidas; el peso a asignar debe estar en relación directa al número de personas de cada distrito. De aquí surge la necesidad de considerar otro estadístico denominado **media ponderada** que se puede considerar como una generalización de la media.

En general, sean x_1, x_2, \dots, x_n las observaciones; por x_i notaremos una observación genérica; la media de ese conjunto de datos, se puede expresar de la siguiente forma

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Hay ocasiones en que no todas las observaciones tienen igual importancia; si w_i es el peso de la observación x_i , la media ponderada se define por el cociente

$$\bar{x}_p = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

Si consideramos que las observaciones x_i son el gasto por habitante en los diferentes distritos y las importancias o pesos son las poblaciones de cada distrito, la expresión anterior queda en la forma

$$\frac{(73423)(12870) + (57110)(12301) + \dots + (74937)(16403)}{73423 + 57110 + \dots + 74937} = 14901$$

que coincide con el valor medio antes calculado. Si a todas las observaciones se les da la misma importancia, es decir, todos los pesos w_i son iguales, la media ponderada coincide con la media aritmética.

Otra medida de tendencia central es la **mediana** que es el valor que, ordenadas las observaciones en forma creciente, ocupa la posición central; en el caso de que tengamos un número n impar de observaciones, la mediana ocuparía el lugar $(n+1)/2$ y en el caso de un número par de observaciones la mediana sería la media de las dos observaciones que ocupan las posiciones centrales. Para el caso del gasto medio por persona, los valores ordenados son

12301, 12870, 14623, 14949, 14957, 15430, 16403, 17105

Como hay un número par de observaciones, los dos valores centrales son 14949 y 14957 y por tanto la mediana sería el valor

$$\frac{14949 + 14957}{2} = 14953$$

Aparte de las anteriores, otras medidas de tendencia central de uso frecuente son los denominados **percentiles** que se pueden entender como generalizaciones de la mediana; así, el percentil i es aquel valor que, ordenadas las observaciones en forma creciente, el $i\%$ de ellas son menores que él y el $(100-i)\%$ restante son mayores; por ejemplo, que la duración del tratamiento aplicado a un paciente está en el percentil 80 del tiempo empleado en el tratamiento de otros pacientes significa que el 80% de los individuos han sido tratados durante un periodo de tiempo menor que el suyo y el 20% restante tuvieron un tiempo de tratamiento más prolongado; a los percentiles 25, 50 y 75 se les denomina **primer, segundo y tercer cuartil**, respectivamente; así, el segundo cuartil coincide con la mediana.

Mediante el polígono de frecuencias acumulativas relativas se puede calcular gráficamente cualquier percentil; si, por ejemplo, queremos calcular los tres cuartiles de la distribución de la duración del tratamiento mediante litotricia extracorpórea en los

200 individuos con litiasis, no tenemos más que trazar por los valores 0.25, 0.5 y 0.75 del eje de ordenadas rectas paralelas al eje de abscisas; cuando intercepten al polígono de frecuencias acumulativas relativas, trazamos una vertical y el valor de la intersección de ésta con el eje de abscisas es el cuartil correspondiente; la Fig. II.10 ilustra este modo de proceder, donde el primer cuartil es 47, la mediana 59.5 y el tercer cuartil 68.

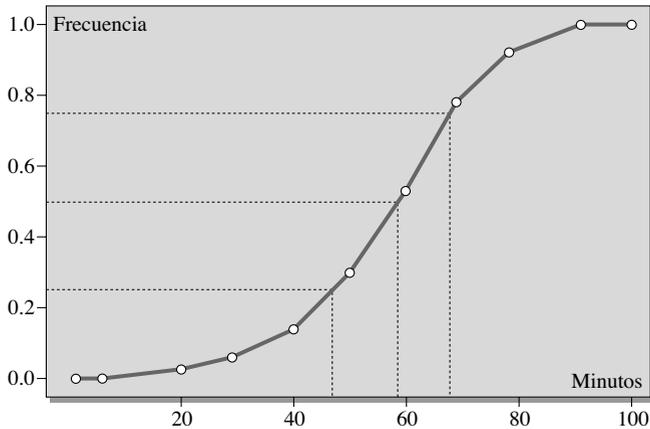


Fig. II.10.- Cálculo gráfico de los tres cuartiles.

Parece que entre las medidas de tendencia central, la mediana es la más fácil de calcular y entender; sin embargo, es la media la que más utilizaremos y la razón principal de ello es, aparte de su mayor popularidad y de su más fácil tratabilidad matemática, lo que posteriormente denominaremos su distribución en el muestreo. Sin embargo, la mediana tiene la ventaja sobre ella de no dejarse influenciar por las **observaciones extremas o raras** de la distribución, propiedad deseable para cualquier medida que intente representar un conjunto de datos. Esta preocupación por controlar el efecto que algunas veces tienen las observaciones extremas, ha llevado a la introducción del concepto estadístico de **robustez** como propiedad consistente en no dejarse afectar demasiado por observaciones raras.

4.2. Medidas de dispersión o variabilidad

Aunque las medidas de tendencia central proporcionan un valor numérico que intenta resumir nuestro conjunto de observaciones, la información que aportan es limitada; en efecto, los dos conjuntos de observaciones

18, 19, 20, 21, 22 y 9, 14, 20, 27, 30

serían indistinguibles a la hora de resumirlos mediante las medidas de tendencia central; la media y la mediana de esas dos series es la misma, 20; entonces, ¿qué hay en esas dos series que las diferencie? En la primera los datos son más parecidos entre sí, son más homogéneos, que en la segunda. Por tanto parece conveniente introducir medidas, estadísticas, que indiquen la variabilidad de las observaciones y que complementen la información aportada por las medidas de tendencia central.

La medida más simple de variabilidad de un conjunto de observaciones es el **rango** entendido como la diferencia entre el valor más grande y el más pequeño de la serie; así, mientras el rango de la primera serie anterior es 4, el de la segunda es 21. Otra forma de definir la variabilidad es mediante la distancia, la diferencia, desde cada observación a un valor central; esta noción nos lleva a la medida llamada **desviación media d_m** :

$$d_m = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

que para la primera serie sería $(1/5)(|18-20|+|19-20|+|20-20|+|21-20|+|22-20|) = 6/5 = 1.2$ mientras que para la segunda es $(1/5)(|9-20|+|14-20|+|20-20|+|27-20|+|30-20|) = 34/5 = 6.8$

La Fig. II.11 muestra ambos conjuntos de observaciones y la distancia de cada valor numérico a la media, representada mediante una línea horizontal de trazo grueso. La primera de las medidas de variabilidad vista anteriormente, el rango, está representada por las líneas punteadas y la desviación media no es más que la suma de las longitudes de los segmentos verticales dividida por 5; gráficamente puede apreciarse la diferencia que existe entre las dos series numéricas teniendo en cuenta sus dispersiones.

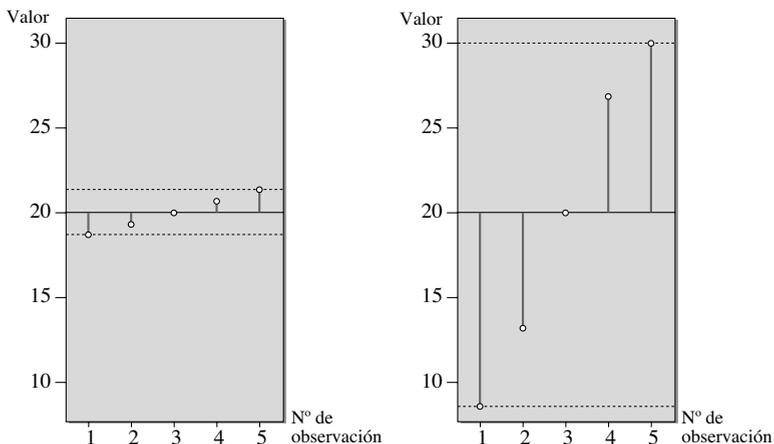


Fig.II. 11.- Representación de las desviaciones respecto de la media.

El indicador de variabilidad más ampliamente usado es la **desviación típica** que es definida como la raíz cuadrada de otra medida llamada **varianza**. La fórmula de la varianza es:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

aunque, a efectos de cálculo, es más conveniente la expresión

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}{n - 1}$$

Por tanto, para el cálculo de la varianza necesitamos conocer la suma de los cuadrados de las observaciones, $\sum_{i=1}^n x_i^2$, y el cuadrado de la suma de las observaciones, $(\sum_{i=1}^n x_i)^2$

Así, para la primera serie anterior, la suma de los cuadrados es:

$$18^2 + 19^2 + 20^2 + 21^2 + 22^2 = 2010$$

y el cuadrado de la suma es

$$(18 + 19 + 20 + 21 + 22)^2 = 10000$$

Por tanto, la varianza de esa serie es:

$$s^2 = \frac{(2010 - \frac{1}{5} 10000)}{4} = 2.5$$

y su desviación típica es $\sqrt{2.5} = 1.58$

De la misma manera, la varianza de la segunda serie es:

$$s^2 = \frac{(2306 - \frac{1}{5} 10000)}{4} = 76.5$$

y su desviación típica es $\sqrt{76.5} = 8.75$

Aunque la varianza y la desviación típica contienen la misma información, quizás la segunda sea más comprensible en cuanto que mientras la varianza se expresa en las unidades de la variable al cuadrado, la desviación típica se expresa en las mismas unidades que la variable medida. Como medida de dispersión se utiliza también el **recorrido intercuartílico**, que no es más que la diferencia entre el tercer y primer cuartil. De las definiciones anteriores se puede deducir que cualquier medida de variabilidad puede ser, como poco cero, es decir, ninguna medida de variabilidad puede tomar valores negativos.

Hay ocasiones en que se está interesado en comparar la heterogeneidad de dos series de datos; si la variable medida es la misma en las dos series, la desviación típica nos puede valer para ese fin; pero imaginemos que en una serie de individuos medimos dos variables, el peso que expresamos en kilogramos y la altura que lo hacemos en metros; en éste caso no podemos comparar desviaciones típicas pues estas vienen expresadas en las mismas unidades que las variables medidas; esta es la razón para definir una variabilidad relativa mediante otra medida denominada **coeficiente de variación** (C.V.) que no es más que el cociente entre la desviación típica y la media de cada serie, medida que ya es adimensional y que no tiene el problema antes aludido;

$$C.V. = \frac{s}{\bar{x}}$$

que se puede expresar en tanto por uno o en forma de porcentaje.

5. Otra forma de resumir una serie de datos

Existe una representación gráfica que, aparte de mostrar la forma de la distribución, recoge también alguna de las medidas que acabamos de definir; se le conoce con el nombre de «**caja**» (*box-plot* o *box and whisker plot*) y se construye de la siguiente forma: sobre el eje en donde se representan los valores de la variable, se construye un rectángulo de anchura cualquiera y altura igual al recorrido intercuartílico; dentro de él se traza un segmento en el punto correspondiente a la mediana; por último, se definen lo que se llaman los **valores adyacentes** de la siguiente manera: el **valor adyacente inferior** es el valor más pequeño de la distribución tal que sea mayor o igual que la cantidad

$$Q_1 - 1.5(Q_3 - Q_1) = O_1$$

donde Q_1 , Q_3 son el primer y tercer cuartil respectivamente y, por tanto, la diferencia $Q_3 - Q_1$ es el recorrido intercuartílico. De la misma forma, el **valor adyacente superior**

es la observación más grande de la serie y tal que sea menor o igual que la cantidad

$$Q_3 + 1.5 (Q_3 - Q_1) = O_2$$

Una vez calculados los valores adyacentes superior e inferior, se trazan segmentos desde las caras laterales del rectángulo hasta los valores adyacentes, con lo que queda completada la caja. Según el criterio de Tukey, los valores de la distribución que sean menores que el valor adyacente inferior o mayores que el valor adyacente superior se denominan **observaciones extremas o outliers**, es decir, observaciones «raras» respecto al conjunto de todas las observaciones.

Como ejemplo, consideremos los datos de la presión sistólica de los 80 adultos; en primer lugar tendríamos que calcular los tres cuartiles para lo que podemos utilizar el tronco y hoja, pero ordenando las hojas para cada tronco, como aparece en la siguiente figura

```

09 | 9
10 | 0 1 4 4 5 5 6 6 6 7 7 8 8 8 9
11 | 0 0 1 1 2 2 2 2 3 3 3 4 5 7 7 7 7 7 8 8 8 8 9 9 9 9 9
12 | 0 0 0 0 0 1 1 1 1 2 2 2 2 2 3 3 3 3 5 6 7 7 7 8 8 8 8 8 9 9
13 | 3 3 5 8 9
14 | 1 2

```

Los tres cuartiles son 111.5, 119 y 123 respectivamente, por lo que los valores adyacentes superior e inferior serían:

- 1) Valor adyacente inferior es el menor valor de la distribución que sea superior o igual a

$$111.5 - 1.5 (123 - 111.5) = 94.25$$

por tanto, el valor adyacente inferior es 99.

- 2) El valor adyacente superior es el mayor valor de la distribución tal que sea menor o igual que

$$123 + 1.5 (123 - 111.5) = 140.25$$

por lo que el valor adyacente superior es 139.

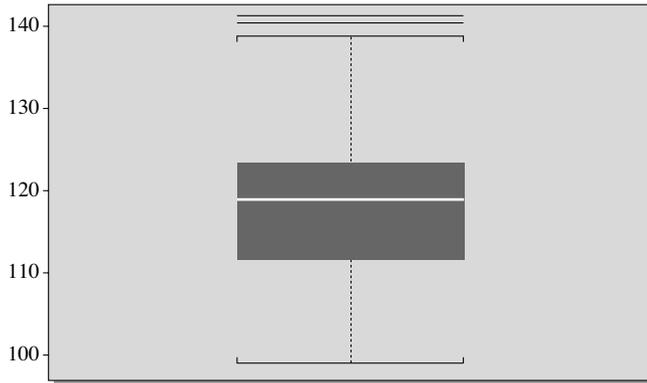


Fig. II. 12.- Representación mediante una caja de las presiones de 80 individuos.

La Fig. II.12 muestra, mediante una caja, la distribución de los 80 individuos; obsérvese cómo a partir de la caja, además de tener idea de tres parámetros de posición central, los tres cuartiles, y de una de variabilidad, el recorrido intercuartílico, nos informa acerca de la posible asimetría de la distribución. Cuanto más centrada esté la mediana en la caja y más similares sean las distancias desde ella hasta los valores adyacentes, más simétrica será la distribución. De acuerdo al criterio antes enunciado sobre las observaciones extremas, y ya que el recorrido intercuartílico para nuestro ejemplo es $123 - 111.5 = 11.5$, las observaciones extremas serán aquellas que caigan fuera del intervalo

$$(111.5 - 1.5 (11.5) , 123 + 1.5(11.5))$$

es decir

$$(94.25 , 140.25)$$

lo que, siguiendo el criterio de Tukey, nos lleva a declarar a los individuos con presiones 141 y 142 como raros, en cuanto a su presión sistólica, respecto al resto de los individuos; estas dos presiones corresponden a los dos segmentos horizontales que quedan por encima del valor adyacente superior.

5.1. Comparación de varias distribuciones

La caja es un instrumento que nos permite la comparación gráfica de distintas distribuciones. Se conoce con bastante certeza que la inactividad física es un factor de riesgo para el infarto de miocardio; a continuación aparece el valor de una medida de riesgo, el riesgo relativo, en diferentes estudios llevados a cabo con el objeto de eva-

luar el riesgo de sufrir infarto entre los individuos que no hacen ejercicio; también se evaluó una medida la calidad del estudio, como aparece en la Tabla II.9, indicando el valor 0 mala calidad, el 1 calidad aceptable y 2 ó 3 buena. Es de interés comparar las distribuciones de los riesgos relativos en las tres categorías de calidad de los estudios.

Tabla II.9.- Riesgo relativo y calidad en 41 estudios sobre inactividad física y enfermedad coronaria.

Riesgo relativo	Calidad del estudio	Riesgo relativo	Calidad del estudio
2.0	0	2.0	1
2.3	0	2.0	1
1.9	0	2.6	2
1.5	1	2.4	2
0.5	0	2.3	1
1.1	0	1.6	1
2.0	1	1.5	1
1.2	2	0.9	0
1.8	0	1.5	0
1.1	0	1.2	1
2.8	1	0.5	0
1.9	1	2.5	0
3.1	2	1.6	2
2.0	1	1.6	2
1.6	1	2.2	0
2.5	1	1.5	0
1.1	0	1.1	0
2.4	0	1.1	0
1.6	2	1.7	0
2.5	3	2.5	3
1.5	0		

Para representar mediante cajas las distribuciones de los riesgos en las tres categorías, tendremos que calcular los tres cuartiles así como los valores adyacentes en cada categoría de calidad. Para calidad mala (0), los riesgos ordenados de forma ascendente son

0.5, 0.5, 0.9, 1.1, 1.1, 1.1, 1.1, 1.1, 1.5, 1.5, 1.5, 1.7, 1.8, 1.9, 2.0, 2.2, 2.3, 2.4, 2.5

por lo que los valores de los tres cuartiles son 1.1, 1.5 y 1.95 aproximadamente. Para conocer los valores adyacentes para esta serie calculamos

$$1.1 - 1.5(1.92 - 1.1) = -0.13$$

y

$$1.92 + 1.5(1.92 - 1.1) = 3.15.$$

por lo que los valores adyacentes serán 0.5 y 2.5. Ya que en la serie no hay ningún riesgo fuera del intervalo (0 , 3.15) no podemos decir que existan riesgos relativos «raros» en ella. Procediendo de igual manera para las dos restantes categorías de calidad del estudio, el lector puede comprobar que los cuartiles que definen las respectivas cajas son: 1.6, 2, 2 para calidad 1 y 1.6, 2.4, 2.5 para calidad 2 o más. Los valores adyacentes para las series son

$$1.2, 2.5 \quad \text{y} \quad 1.2, 3.1$$

por lo que el valor 2.8 de la segunda serie puede considerarse como una observación extrema; en definitiva, las tres distribuciones se pueden representar como aparece en la Fig. II.13.

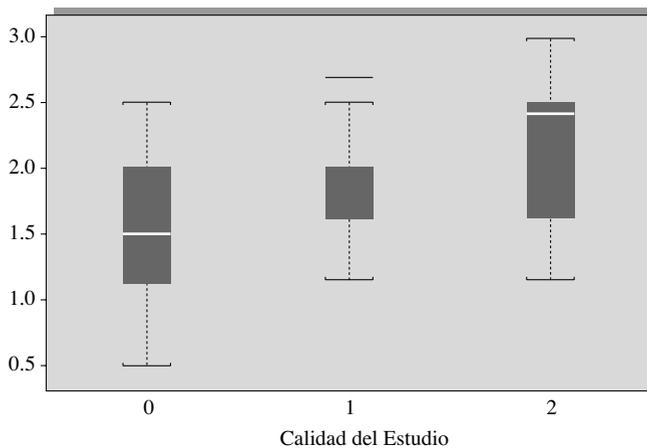


Fig. II.13.- Distribución de los riesgos relativos según la calidad del estudio.

A partir de esta representación gráfica podemos ver cómo a medida que la calidad de los estudios es mayor la evidencia de que la inactividad física es un factor de riesgo para el infarto de miocardio se hace más palpable; entre los estudios de mala calidad hay algunos, aquellos que muestran riesgos relativos inferiores a la unidad, que aseguran que la inactividad física protege del infarto de miocardio.

B. TEST DIAGNOSTICOS

J.A. Chicharro Molero

R. Burgos Rodríguez

1. Introducción

En la práctica médica diaria, los profesionales suelen preguntarse cuestiones sobre el diagnóstico que, en ocasiones, son difíciles de contestar. Algunas de ellas pueden ser:

- 1) ¿La interpretación de la anamnesis y la exploración física es correcta?
- 2) ¿La nueva información sobre el diagnóstico de esta enfermedad es reproducible en este caso?
- 3) ¿Qué exploraciones complementarias nos darán más y mejor información sobre el diagnóstico?

En primer lugar hemos de tener en cuenta que los datos obtenidos de la anamnesis y la exploración física no revelan, usualmente, el verdadero estado del paciente. Excluyendo los fallos metodológicos, este hecho se debe a que los signos y síntomas pueden ser representativos de más de una enfermedad:

Un varón de 60 años con un síndrome miccional intenso de carácter irritativo acude a la consulta, no presentando fiebre ni hematuria microscópica.

Este cuadro podría estar producido por entidades tan diferentes como una infección del tracto urinario bajo, agudización de una hiperplasia benigna prostática o una cáncer in situ de vejiga. En cualquier caso un tacto rectal, analítica de orina y, si procediera, una citología urinaria nos sacarían de dudas; pero no con el 100%, ya que estas pruebas no son exactas y tienen un porcentaje de fallos como cualquier otra, ya que incluso los estudios histopatológicos son, en ocasiones, difíciles de interpretar y varían con la pericia del patólogo. Mediante esta breve reflexión podemos extraer dos conclusiones: una, que no sólo un único test diagnóstico nos dará la certeza diagnós-

tica, dos que nunca ésta será del 100%, existiendo siempre incertidumbre sobre si la prueba complementaria acertó o no en el diagnóstico.

Esta incertidumbre, inherente a la práctica médica, se suele caracterizar por palabras como «probablemente o posiblemente» que son bastante imprecisas ya que no cuantifican con exactitud el fallo que podemos esperar al aplicar un test diagnóstico.

Un varón de 25 años es diagnosticado de masa testicular que resulta ser un tumor de células germinales no seminomatoso. En el estudio de extensión no se evidencia afectación ganglionar ni metástasis, siendo los marcadores negativos antes y después de la orquiectomía, tras ella el urólogo le plantea al paciente la necesidad de linfadenectomía, ante lo cual éste demanda por qué y el urólogo le contesta que «no se puede excluir la diseminación a los ganglios», alarmado el paciente requiere una segunda opinión a otro urólogo, éste le confirma la necesidad de linfadenectomía, pero le dice que «posiblemente no tenga diseminación ganglionar». Ante esto el paciente decide no someterse a la realización de linfadenectomía.

En este caso los dos urólogos tenían razón, pero las palabras ambiguas utilizadas produjeron distintas reacciones en el paciente. En realidad el infraestadiaje clínico de los tumores testiculares es del 20-30% según las series; en consecuencia, si la información sobre la incertidumbre diagnóstica se hubiera cuantificado, el paciente no se hubiese angustiado y hubiese aceptado la linfadenectomía.

Esta gran confusión que se puede originar al cuantificar la incertidumbre con palabras imprecisas, se puede evitar hablando en términos de probabilidad que oscila en valores de 0 a 1, aunque en medicina, sobre todo en la comunicación con los pacientes, se puede expresar en porcentajes. Así, si en el caso anterior los dos urólogos le hubiesen dicho al paciente que «la probabilidad que tenga afectación ganglionar es del 20-30%» se hubieran evitado las ambigüedades que se produce por:

- Los médicos eligen diferentes palabras para expresar un mismo juicio sobre la probabilidad de un suceso futuro.
- Los médicos eligen las mismas palabras para expresar muy diferentes juicios sobre la probabilidad de un proceso.
- Las palabras no pueden describir precisamente cuánto cambiará la opinión sobre una hipótesis diagnóstica, conforme se vayan adquiriendo nuevas informaciones.

Con respecto a las dos últimas preguntas, es totalmente necesario conocer la metodología epidemiológica y estadística para la evaluación de los test diagnósticos, para

así saber interpretar críticamente la información que proporcionan los artículos científicos sobre nuevas pruebas diagnósticas, dilucidando si aportan más valía diagnóstica que los ya existentes y si contuvieran fallos metodológicos que invaliden los resultados de las publicaciones, ya que es tremendamente serio la asunción de conclusiones basadas en estudios que asienten sobre diseños epidemiológicos erróneos y cálculos estadísticos inadecuados.

2. Probabilidad

Existen dos tipos de fenómenos que pueden ocurrir; los **deterministas** que son los que realizados en las mismas condiciones, siempre dan los mismos resultados; y los **aleatorios** que son aquellos en los que no se puede predecir su resultado aunque se realicen en idénticas situaciones. Desgraciadamente en la práctica médica abundan los segundos y apenas existen, si es que los hay, los primeros. La cuantificación de los fenómenos aleatorios se realiza a través de la **teoría de la probabilidad**.

Probabilidad se puede definir como un valor que oscila de 0 a 1 y que representa la frecuencia relativa con la que se produce un suceso aleatorio determinado cuando la experiencia se repite un número infinito de veces⁷. Es decir, si lanzamos un dado un número determinado de veces, por ejemplo 20, puede ocurrir que el número 5 salga seis veces, el 4 dos, el 3 cinco, el uno tres, el 6 cuatro y el 4 ninguna, pero por la **ley del azar** si repetimos la experiencia un número muy grande de veces, la frecuencia de aparición de cada número se irá estabilizando alrededor de 1/6 que es exactamente la probabilidad que salga un determinado número cada vez que se tira el dado.

Evidentemente cada vez que se lanza el dado es imposible que salgan dos números a la vez, por lo que a este tipo de sucesos aleatorios se les denomina **incompatibles**; así mismo existen sucesos aleatorios en los que la aparición de uno no vincula para la del otro, por lo que se llaman **independientes**⁸; por ejemplo la probabilidad de tener un hipernefroma en el riñón derecho es independiente de tener una litiasis en el riñón izquierdo. Los sucesos independientes nunca pueden ser incompatibles, pues en definitiva estos últimos tiene relación ya que si aparece uno no puede hacerlo el otro simultáneamente.

3. Probabilidad condicionada y teorema de Bayes

En medicina lo más importante y lo que se utiliza cotidianamente en el razonamiento, muchas veces sin tener conciencia de ello, es la llamada **probabilidad condicionada** que se define como la probabilidad de que ocurra un suceso aleatorio una

vez que ha ocurrido otro previo; por ejemplo «¿cuál es la probabilidad de que un paciente tenga un carcinoma de próstata una vez que el PSA es mayor de 10 ng/ml?» que no es lo mismo que plantearse «¿Cual es la probabilidad de que en pacientes con carcinoma de próstata el PSA sea mayor de 10 ng/ml?» . En este punto, para no aburrir al lector no neófito y para despertar al neófito, adelantaremos que ambas preguntas son probabilidades condicionadas; la primera, corresponde al valor predictivo positivo, y la segunda a la sensibilidad.

Continuando de una manera, a nuestro criterio, más pedagógica, la evaluación de test diagnósticos se determina mediante probabilidades condicionadas que se pueden dividir en dos grandes grupos:

- 1) Las que determinan el comportamiento de un test diagnóstico, tacto rectal o PSA, en los sujetos enfermos, con carcinoma de próstata, y en los no enfermos, sin carcinoma de próstata. Son intrínsecas al test diagnóstico y miden su capacidad discriminativa diagnóstica. Son la **sensibilidad** y la **especificidad** y, sus complementarios, la proporción de **falsos negativos** y **falsos positivos** respectivamente.
- 2) Las de un mayor interés en la clínica diaria; es decir cual es la probabilidad de que un sujeto que acude a la consulta tenga una determinada enfermedad, carcinoma de próstata, una vez que le practicamos algún test, tacto rectal y/o PSA. Son los **valores predictivos, positivo y negativo**.

Lo primero que hay que conocer de una prueba diagnóstica es su poder de discriminación, es decir sensibilidad, especificidad y sus complementarios; para ello debemos aplicar el test que queramos analizar (PSA, tacto rectal, ecografía) a un grupo, la muestra, de individuos en el que existan personas con la enfermedad (cáncer de próstata) y personas sin ella, lo cual se debe haber determinado previamente por uno o varios tests diagnósticos que estén considerados como los de máxima fiabilidad (**estándar ideal**). La selección de la muestra debe seguir ciertas normas importantes:

- 1) Debe tener un tamaño adecuado que garantice el control del error aleatorio (significación y potencia).
- 2) El grupo de pacientes control, es decir los que no tienen la enfermedad deben pertenecer a sectores que planteen diagnósticos diferenciales. Es decir si estudiamos el poder diagnóstico del PSA en el cáncer prostático el grupo control no podrá ser varones jóvenes o mujeres; sino que deberán ser pacientes con edades similares e hipertrofia prostática en los que haya sido excluido el cáncer de próstata mediante el estándar ideal.

- 3) Los pacientes que pertenezcan al grupo de enfermos deben representar a todos los estadios de la enfermedad, a no ser que el estudio verse sobre un estadio específico.
- 4) La prevalencia (proporción de pacientes enfermos sobre el total) del grupo de estudio debe de ser lo más aproximada a la prevalencia de la enfermedad en la población ya que como veremos la prevalencia en el grupo de estudio influye grandemente en los estimadores de la utilidad diagnóstica.

Supongamos que deseamos conocer la capacidad discriminativa del PSA en el diagnóstico del cáncer prostático; para ello aplicamos a un grupo de pacientes con y sin la enfermedad el test (PSA), determinando el valor de 10ng/ml como el punto discriminante como normal y patológico, es decir valores mayores nos indicarían presencia de la enfermedad y valores menores indicarían ausencia de ella; posteriormente expresamos los resultados en la siguiente tabla 2x2:

Tabla II.10. Presencia de carcinoma de próstata según nivel de PSA.

		PSA	
		> 10ng/ml	< 10ng/ml
CARCINOMA DE PRÓSTATA	SI	A	B
	NO	C	D

Tras la aplicación del test pueden ocurrir cuatro circunstancias representadas por las casillas:

- Individuos en los que el test clasificó como enfermos, $PSA > 10$, y que realmente lo estaban; son los **verdaderos positivos (A)**.
- Individuos en los que el test clasificó como no enfermos, $PSA < 10$, y que realmente no lo estaban; son los **verdaderos negativos (D)**.
- Individuos en los que el test clasificó como enfermos, pero que realmente no estaban enfermos, es decir serán los **falsos positivos (C)**.
- Individuos clasificados como no enfermos por el test diagnóstico, pero que realmente estaban enfermos, es decir los **falsos negativos (B)**.

La sensibilidad se define como la probabilidad de que el test sea positivo entre los realmente enfermos, $P(\mathbf{T+}/\mathbf{E})$. Dicho de otra manera, la sensibilidad sería la proporción de pacientes realmente enfermos (A+B) con el test positivo, ateniendonos a la tabla sería:

$$\frac{A}{A + B}$$

La especificidad se define como la probabilidad de que el test sea negativo entre los realmente no enfermos, $P(\mathbf{T-}/\mathbf{noE})$. Dicho de otra manera, la especificidad sería la proporción de pacientes realmente no enfermos (C+D) con el test negativo, ateniendonos a la tabla sería :

$$\frac{D}{C + D}$$

La proporción de falsos negativos (**PFN**) se define como la probabilidad de que el test sea negativo entre los realmente enfermos, $P(\mathbf{T-}/\mathbf{E})$; es la probabilidad complementaria a la Sensibilidad y su valor será siempre 1-Sensibilidad. Dicho de otra manera sería la proporción de pacientes realmente enfermos (A+B) con el test negativo, ateniendonos a la tabla sería :

$$\frac{B}{A + B}$$

La proporción de falsos positivos (**FPF**) se define como la probabilidad de que el test sea positivo entre los realmente no enfermos, $P(\mathbf{T+}/\mathbf{noE})$; es la probabilidad complementaria a la Especificidad y su valor será siempre 1-Especificidad. Dicho de otra manera sería la proporción de pacientes realmente no enfermos (C+D) con el test positivo, ateniendonos a la tabla sería:

$$\frac{C}{C + D}$$

Todo este desarrollo es perfecto para ver la capacidad discriminativa de los test diagnósticos, pero en la práctica los urólogos no sabemos a priori, que sujetos tienen cáncer y cuales no y lo que nos interesa en la realización de nuestra actividad asistencial es conocer cuan de seguro estamos de que un sujeto tenga o no la enfermedad (cáncer de próstata) en función de unos test que apliquemos. Pues bien cuando un varón acude a la consulta la probabilidad de que padezca un carcinoma de próstata,

en ausencia de información previa, será igual a la prevalencia de esta enfermedad en la población masculina, a esta probabilidad se le llama **probabilidad a priori** o **probabilidad pretest**. En general la probabilidad de que una persona cualquiera pueda tener una enfermedad determinada, sin ninguna información previa, es igual a la prevalencia de esa enfermedad y la notación matemática es $P(E)$. Una vez que tengamos información sobre el interrogatorio y la exploración física la probabilidad de tener el carcinoma de próstata cambiará dependiendo de los resultados, es decir, estamos ante una probabilidad condicionada que se define como **probabilidad posterior** o **probabilidad posttest**. En general estamos ante la probabilidad de que un paciente tenga la enfermedad (E) una vez que el test ha dado positivo (**T+**), circunstancia que matemáticamente se anota como $P(E/T+)$ que a simple vista pudiera creerse que es la sensibilidad, pero esto es erróneo.

El desarrollo matemático del **Teorema de Bayes**^{9,10} se debe a un clérigo inglés, con el mismo apellido y de nombre Thomas, que vivió en el siglo XVIII (1702-1761), y permite calcular la probabilidad posterior o posttest, por ejemplo la probabilidad de tener la enfermedad (cáncer de próstata) una vez que el test es positivo ($PSA > 10$ ng/ml), $P(E/T+)$; o la probabilidad de no tener la enfermedad cuando el test es negativo, $P(\text{no}E/T-)$, conociendo tres probabilidades:

- 1) La probabilidad anterior que es la prevalencia si no se aplicó otro test previamente a la determinación del PSA; si esto no fué así, será la prevalencia en la muestra del estudio.
- 2) La probabilidad de que el test dé positivo, en nuestro ejemplo $PSA > 10$ ng/ml, cuando se aplica a un grupo de pacientes con la enfermedad, carcinoma de próstata. Esta probabilidad condicionada se anota como $P(T+/E)$ y que como hemos visto se denomina **proporción de verdaderos positivos** o sensibilidad. También puede aparecer la probabilidad complementaria (1-Sensibilidad) que se denomina **proporción de falsos negativos** $P(T-/E)$.
- 3) La probabilidad de que el test dé negativo, $PSA < 10$ ng/ml, cuando se aplica a un grupo de pacientes que no tienen la enfermedad. Se anota como $P(T-/noE)$ y se denomina **proporción de verdaderos negativos** o especificidad. Así mismo, puede aparecer la probabilidad complementaria (1-Especificidad) que se denomina **proporción de falsos positivos** $P(T+/noE)$.

Una vez en el desarrollo del teorema de Bayes, supongamos que queremos calcular la probabilidad de tener carcinoma de próstata cuando tenemos un $PSA > 10$ ng/ml, $P(E/T+)$, por la ley de la probabilidad condicionada sabemos que:

$$P(E/T+) = \frac{P(T+ \text{ y } E)}{P(T+)} \quad (1)$$

El denominador de esta extraña ecuación, se define como la probabilidad de que el test sea positivo, PSA>10ng/ml. Esto será igual a la probabilidad de que el test dé positivo y los pacientes estén enfermos más la probabilidad de que el test sea positivo y los pacientes no estén enfermos, con y sin carcinoma de próstata, al ser sucesos incompatibles, matemáticamente sería:

$$P(T+) = P(T+ \text{ y } E) + P(T+ \text{ y } \text{no}E) \quad (2)$$

Por lo que la fórmula original (1) quedaría:

$$P(E/T+) = \frac{P(T+ \text{ y } E)}{P(T+ \text{ y } E) + P(T+ \text{ y } \text{no}E)} \quad (3)$$

Tomando otra vez la ley de la probabilidad condicionada, obtendremos las probabilidades de lo que llamamos sensibilidad $P(T+/E)$ y proporción de falsos positivos $P(T+/\text{no } E)$:

$$P(T+/E) = \frac{P(T+ \text{ y } E)}{P(E)} \quad P(T+/\text{no } E) = \frac{P(T+ \text{ y } \text{no}E)}{P(\text{no}E)}$$

Despejando en ambas fórmulas obtenemos:

$$P(T+ \text{ y } E) = P(E) * P(T+/E)$$

$$P(T+ \text{ y } \text{no}E) = P(\text{no}E) * P(T+/\text{no}E)$$

Sustituyendo los valores en la fórmula número (3) obtendremos el teorema de Bayes:

$$P(E/T+) = \frac{P(E) * P(T+/E)}{P(E) * P(T+/E) + P(\text{no}E) * P(T+/\text{no}E)} \quad (4)$$

Pues bien, la probabilidad hallada $P(E/T+)$ y que se enuncia como la probabilidad de que un paciente tenga la enfermedad, carcinoma de próstata, una vez que el

test es positivo, PSA>10ng/ml, se denomina **valor predictivo positivo (VPP)** y depende como ya se dijo anteriormente de la sensibilidad, P(T+/E), proporción de falsos positivos, P(T+/noE), y prevalencia de la enfermedad en el grupo de estudio P(E), que como veremos es importantísima para los valores predictivos. Es decir, la fórmula (4) traducida del sánscrito quedaría de la siguiente forma:

$$VPP = \frac{\text{Sensibilidad} * \text{Prevalencia}}{\text{Sensibilidad} * \text{Prevalencia} + \text{PFP} * (1-\text{Prevalencia})} \quad (4b)$$

Asimismo es de interés saber la probabilidad de que un paciente no esté enfermo, no tenga carcinoma de prostata, cuando el test da negativo, PSA<10ng/ml; esta probabilidad se denomina **valor predictivo negativo (VPN)** y, según una deducción análoga a la anterior, siguiendo el teorema de Bayes es:

$$P(\text{noE}/T-) = \frac{P(\text{noE}) * P(T-/noE)}{P(\text{noE}) * P(T-/noE) + P(E) * P(T-/E)} \quad (5)$$

Que traducida quedaría:

$$VPN = \frac{\text{Especificidad} * (1-\text{Prevalencia})}{\text{Especificidad} * (1-\text{Prevalencia}) + \text{PFN} * \text{Prevalencia}} \quad (5b)$$

Observando las fórmulas de ambos valores predictivos, podemos apreciar que los dos dependen de tres probabilidades que son:

- La especificidad P(T-/noE), o su complementario, la proporción de falsos positivos P(T+/noE).
- La sensibilidad P(T+/E), o su complementario, la proporción de falsos negativos P(T-/E).
- La prevalencia P(E), o su complementario, la proporción de no enfermos P(noE).

Para el sufrido lector, que se maree con las fórmulas, diremos que mediante artificios matemáticos, el cálculo de los valores predictivos se pueden hacer desde la tabla 2x2; así el VPP sería A/A+C, y el VPN sería D/B+D. Aunque, evidentemente, el cálculo desde la tabla 2x2 es más rápido y fácil, personalmente aconsejamos las fórmulas del Teorema de Bayes, pues como veremos con posterioridad son de ayuda incalculable para la comprensión de la influencia de sensibilidad, especificidad y prevalencia en los valores predictivos.

4. Influencia de sensibilidad, especificidad y prevalencia en valores predictivos

Como ya hemos comentado la Sensibilidad y Especificidad son probabilidades intrínsecas a un test diagnóstico determinado; por otra parte también hemos visto que la prevalencia es una probabilidad pretest. Teniendo en cuenta que los valores predictivos son probabilidades que se obtienen tras la aplicación de un test diagnóstico (postest), es fácil comprender que dependerán de la probabilidad pretest y de las probabilidades intrínsecas del test a aplicar; por otra parte si el abnegado lector ha tenido en cuenta nuestro consejo del párrafo anterior verá como estas tres últimas probabilidades están inmersas en las fórmulas que proporciona el teorema de Bayes para los valores predictivos.

Hasta ahora sabemos que ambos valores predictivos dependen de tres probabilidades, pero no dependen por igual de ellas.

4.1. Valor Predictivo Positivo

- a) *Sensibilidad/Especificidad.*- Si la proporción de falsos positivos de un test es muy baja querrá decir que la mayoría de los sujetos con test positivo estarán en el grupo de enfermos, por lo que cuando el test sea positivo, será bastante probable encontrar enfermos; es decir $P(E/T+)$ será grande. De una forma más gráfica, observando la fórmula 4b del teorema de Bayes, se aprecia que si la proporción de falsos positivos (FPF) es 0, la fórmula quedaría como sigue:

$$VPP = \frac{\text{Sensibilidad} * \text{Prevalencia}}{\text{Sensibilidad} * \text{Prevalencia} + 0}$$

Con esta fórmula, independientemente de los valores para Sensibilidad y Prevalencia, el VPP valdrá 1, que en términos de probabilidad es el mayor que puede tener. Teniendo en cuenta que FPF es igual que 1-Especificidad, ya que son complementarios, en este supuesto la Especificidad tendrá su valor máximo, 1. Por lo que *cuando la Especificidad sea 1 el VPP también valdrá 1*. Con la Sensibilidad no pasa este hecho ya que se encuentra en numerador y denominador.

- b) *Prevalencia.* Es lógico pensar que si la prevalencia es alta, la proporción de enfermos es muy alta y cuando el test sea positivo la probabilidad de ser enfermo $P(E/T+)$ será muy alta. Observando de nuevo la fórmula 4b del teorema de

Bayes, podemos apreciar que para prevalencias de uno, su valor máximo, el segundo sumando del denominador valdrá cero quedando la fórmula:

$$VPP = \frac{\text{Sensibilidad} * \text{Prevalencia}}{\text{Sensibilidad} * \text{Prevalencia} + 0}$$

Y el VPP valdría 1. Sin embargo cuando la proporción de enfermos sea pequeña, prevalencia baja, una vez que el test sea positivo será más probable encontrar no enfermos, aumentará $P(\text{noE}/\text{T}+)$, y por tanto disminuirá el VPP, $P(\text{E}/\text{T}+)$, que es su complementaria. Conforme aumenta la prevalencia aumenta el VPP y viceversa

4.2. Valor Predictivo Negativo

- a) *Sensibilidad/Especificidad*. Si la proporción de falsos negativos de un test es muy baja querrá decir que la mayoría de los sujetos con test negativo estarán en el grupo de no enfermos, por lo que cuando el test sea negativo, será bastante probable encontrar no enfermos; es decir $P(\text{noE}/\text{T}-)$ será grande. De una forma más gráfica, observando la fórmula 5b del teorema de Bayes, se aprecia que si la proporción de falsos negativos (PFN) es 0, la fórmula quedaría como sigue:

$$VPN = \frac{\text{Especificidad} * (1 - \text{Prevalencia})}{\text{Especificidad} * (1 - \text{Prevalencia}) + 0}$$

Con esta fórmula, independientemente de los valores para Especificidad y Prevalencia, el VPN valdrá 1, que en términos de probabilidad es el mayor que puede tener. Teniendo en cuenta que PFN es igual que $1 - \text{Sensibilidad}$, ya que son complementarios, en este supuesto la Sensibilidad tendrá su valor máximo, 1. Por lo que *cuando la Sensibilidad sea 1 el VPN también valdrá 1*. Con la Especificidad no pasa este hecho ya que se encuentra en numerador y denominador.

- b) *Prevalencia*. Es lógico pensar que si la prevalencia es baja, la proporción de no enfermos es muy alta y cuando el test sea negativo la probabilidad de no ser enfermo $P(\text{noE}/\text{T}-)$ será muy alta. Observando de nuevo la fórmula 5b del teorema de Bayes, podemos apreciar que para prevalencias de cero, su valor mínimo, el segundo sumando del denominador valdrá cero quedando la fórmula:

$$\text{VPN} = \frac{\text{Especificidad} * (1 - \text{Prevalencia})}{\text{Especificidad} * (1 - \text{Prevalencia}) + 0}$$

Y el VPN valdría 1. Sin embargo cuando la proporción de enfermos sea grande, prevalencia alta, una vez que el test sea negativo será más probable encontrar enfermos, aumentará P(E/T-), y por tanto disminuirá el VPN, P(noE/T-), que es su complementaria. Conforme disminuye la prevalencia aumenta el VPN y viceversa.

En resumen, para obtener valores predictivos positivos altos, necesitaremos test muy específicos y para valores predictivos negativos altos necesitaremos test muy sensibles. Pero, ojo con la prevalencia ya que a valores altos, aumenta el VPP y disminuye el VPN, y viceversa; en este momento se nos ocurre una pregunta para la distensión del sufrido lector que haya osado llegar hasta aquí, ¿para que sirve todo este esfuerzo metodológico, si todos los pacientes tuvieran la enfermedad?, la respuesta es obvia, para nada en absoluto.

De todo lo dicho hasta ahora sobre la influencia en los valores predictivos se puede llegar a varias conclusiones:

- 1) Vigilar el número de pacientes con la enfermedad en el grupo de estudio (prevalencia) para interpretar bien los datos.
- 2) La importancia de la proporción de falsos positivos (1-especificidad) en el valor predictivo positivo es tanto menor cuanto mayor es la prevalencia, ya que un VPP alto tras la aplicación de un test será importante si aporta mucha información sobre la que teníamos antes de la aplicación del mismo.
- 3) La importancia de la proporción de falsos negativos (1-sensibilidad) en el valor predictivo negativo es tanto menor cuanto menor es la prevalencia, ya que un VPN bajo tras la aplicación de un test será importante si aporta mucha información sobre la que teníamos antes de la aplicación del mismo.
- 4) Para evitar el problema que puede suponer que la muestra no sea representativa de la población de interés, bien por que haya demasiados pacientes enfermos o no enfermos, podemos obtener la **razón de probabilidades**^{9, 11} que es el cociente entre la sensibilidad y la proporción de falsos positivos. Este valor apenas se influencia por una composición de la muestra diferente a la población de interés.

5. Test diagnósticos y variables continuas

5.1. Conceptos Generales

Hasta ahora hemos visto test diagnósticos cuyos resultados se miden como variables categóricas, por ejemplo tacto rectal, y cuando hemos tratado variables continuas como el PSA estableciendo un valor, artificialmente claro, que determinaba el estado de enfermo y no enfermo. En la realidad la total independencia de la distribución de una variable (glucemia, PSA, etc.) en enfermos y no enfermos, situación (a) de la Fig. II.14., ocurre muy pocas veces, si es que hay alguna, ya que existe solapamiento de ambas distribuciones, situación (b) de la Fig. II.14, tal que si establecemos el punto de corte que discrimine lo normal de lo patológico en un valor determinado, por ejemplo $PSA > 7$, habrá enfermos de cáncer de próstata con valores menores (falsos negativos) y no enfermos, sin cáncer de próstata, con valores mayores (falsos positivos).

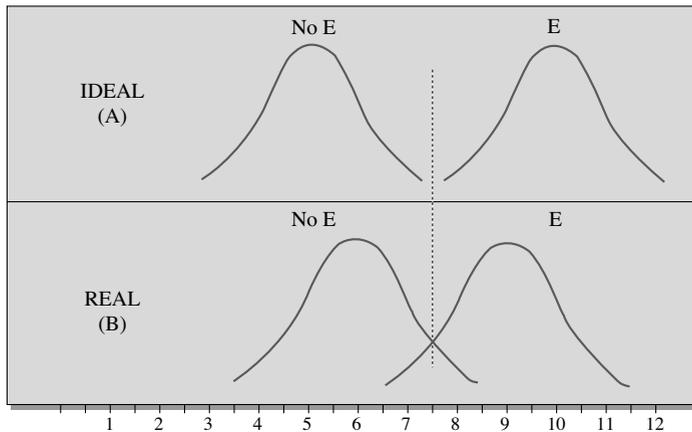


Figura II.14. Variables cuantitativas y punto de corte.

Si estableciéramos diferentes puntos de corte para el PSA en un mismo grupo se podría observar cómo cambian los valores de las casillas de la tabla 2X2, ya que los falsos positivos (C), falsos negativos (B), verdaderos positivos (A) y verdaderos negativos (D) cambian, permaneciendo el número de enfermos (A+B) y no enfermos invariable, siendo, por tanto, la sensibilidad y la especificidad diferentes para cada punto de corte. Este hecho tendrá una enorme trascendencia, ya que si el nivel normal de PSA lo ciframos en 10ng/ml obtendremos una sensibilidad y especificidad determinadas, pero si el valor que discrimina lo normal y lo patológico lo descendemos a 4, ¿qué pasará?; como es lógico pensar, diagnosticaremos más pacientes como patológicos y se nos escaparán menos carcinomas de prostata, con lo que disminuirémos los falsos negativos, y aumentaremos la sensibilidad, pero por el contrario aumentaremos los

falsos positivos y disminuirémos la especificidad (ver Fig. II.15 (a)); en cambio al aumentar el punto de corte, a 10 ng/ml, figura II.15 (b), los falsos negativos serán mayores ya que se nos escapan más carcinomas, disminuyendo los falsos positivos. En resumen conforme aumentamos el punto de corte disminuirá la sensibilidad y aumentará la especificidad y viceversa.

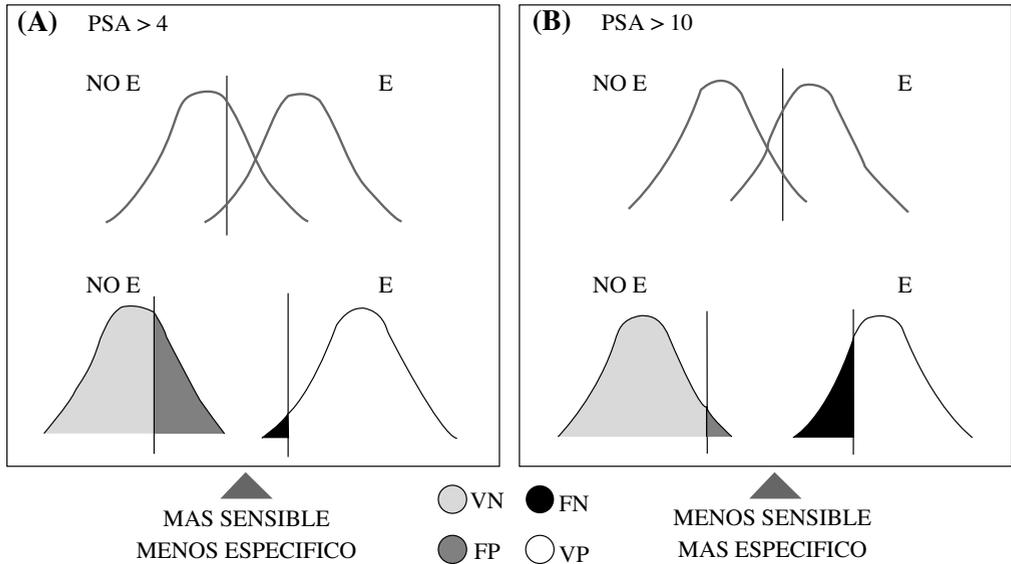


Fig. II.15. Influencia del punto de corte.

Después de todo lo dicho, surge la gran pregunta: *¿cuál es el punto de corte que mejor discriminación diagnóstica tiene?*; pues bien, para contestar a este interrogante debemos recurrir a las **curvas de rendimiento diagnóstico o Curvas ROC**¹¹.

5.2. Curvas ROC

Estas curvas tienen su origen en la tecnología de detección de señales, y más concretamente en la predicción de la trayectoria de misiles¹², siendo posteriormente incorporadas a la evaluación de test diagnósticos. Consisten en la representación gráfica mediante un eje cartesiano (X,Y) de la sensibilidad en el eje Y, y la proporción de falsos positivos (1-especificidad) en el eje X, de cada uno de los puntos de corte, tipo III, existiendo otros tres tipos de curva I,II,IV que representan sensibilidad, especificidad, o los complementarios de éstas, sin embargo nosotros nos referiremos a las ROC tipo III.

La representación gráfica de la curva ROC que aparece en Fig. II.16

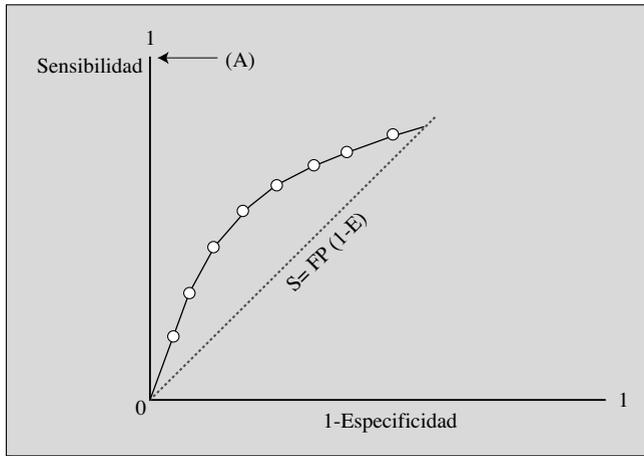


Fig. II. 16. Curva ROC.

- 1) La línea que se sitúa a 45 grados del eje de las X y del eje de las Y, llamada línea de no discriminación. Todos los puntos que se sitúen en esta línea tendrán igual proporción de verdaderos positivos (sensibilidad) que de falsos positivos; es decir el test se comportará de igual manera en el grupo de enfermos que en el de no enfermos, no teniendo, por tanto, ninguna capacidad de discriminación diagnóstica.
- 2) El área por encima de la línea de 45 grados en la que la proporción de probabilidades es mayor de 1, lo que indica que un resultado del test positivo, $PSA > 10\text{ng/ml}$, se dará con más probabilidad en los enfermos que en los no enfermos.
- 3) El área por debajo de la línea de 45 grados; en ésta la proporción de probabilidades es menor de 1, indicando que el test dará positivo con más probabilidad en los no enfermos (falsos positivos) que en los enfermos (verdaderos positivos).
- 4) El punto de máxima discriminación diagnóstica (a) es el situado en el ángulo superior izquierdo, donde sensibilidad y especificidad alcanzan su valor máximo, 1.

Las curvas ROC tienen dos aplicaciones concretas en la evaluación de test diagnósticos cuyos valores sean variables continuas: la elección del punto de corte de máxima discriminación diagnóstica y la comparación de dos test diagnósticos ¹³.

- 1) *Punto de corte óptimo*: Como ya hemos comentado el punto de corte óptimo del test será el que más se aproxime al ángulo superior izquierdo, que coincidirá con la intersección de la curva con la línea trazada perpendicular a la línea de no discriminación (bisectriz) desde el punto donde sensibilidad y especificidad valen 1, como se muestra en la Fig. II.17.

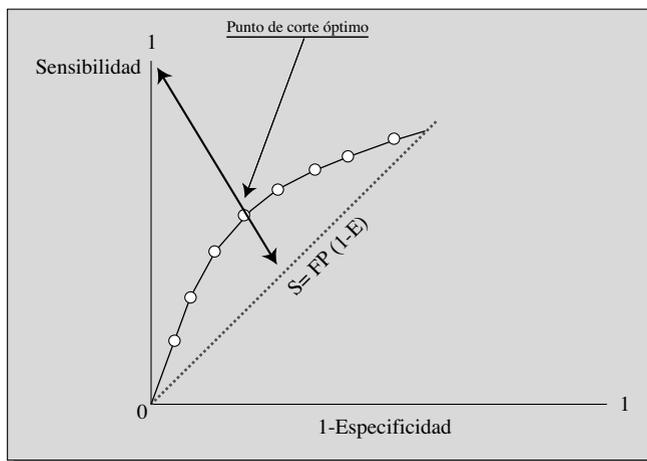


Fig. II. 17. Curva ROC. Punto de corte óptimo.

Una segunda forma de elegir el punto de corte óptimo sería aquel en que $(\text{sensibilidad} + \text{especificidad})/2$ tenga el valor más alto.

- 2) *Comparación de tests diagnósticos*: Como ya hemos comentado, la discriminación de un test viene dada por el área comprendida entre la curva y la línea de no discriminación. Pues bien para efectuar la comparación se representan en el mismo eje las curvas de los test a comparar y aquel con mayor área bajo la curva será el que posea más poder diagnóstico, ver Fig. II.18.

6. Aplicación de varios test diagnósticos

6.1. Tests secuenciales

Es usual en la práctica médica utilizar test diagnósticos una vez que ha sido utilizado otro previo para seleccionar un subgrupo de pacientes que padezcan la enfermedad de interés. Por ejemplo, una vez que hemos utilizado la determinación de PSA, obtenemos un grupo de pacientes cuyos niveles son mayores de 10 ng/ml, pues bien,

a estos pacientes les realizamos biopsia prostática ecodirigida para confirmar aún más el diagnóstico, pues en este grupo habrá individuos que a pesar de tener elevados los niveles de PSA, no tengan carcinoma de prostata. Pero en la aplicación secuencial hemos de tener en cuenta importantes medidas metodológicas.

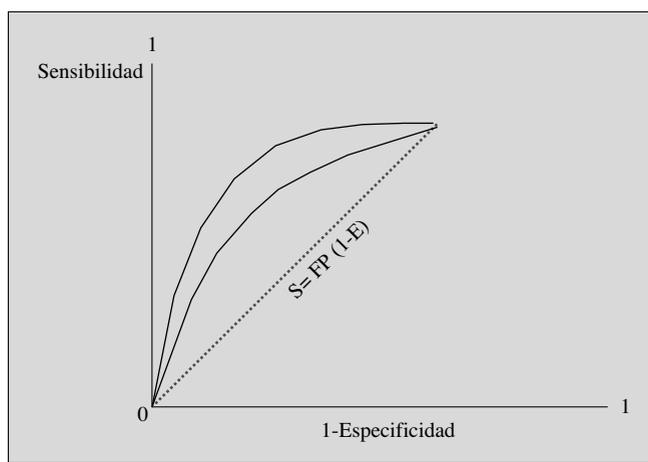


Fig. II.18. Comparación de tests diagnósticos.

En primer lugar que el subgrupo de pacientes que será sometido al segundo test debe ser el determinado como enfermo por el primer test, por lo que será tremendamente importante que la proporción de falsos negativos sea muy baja en el primer test. Así si los valores del PSA entre 4 y 10 mg/ml, nos ofrecen dudas diagnósticas importantes, deberíamos situar el valor discriminante de lo normal y patológico en 4 ng/ml para evitar la pérdida de pacientes que tengan la enfermedad. Con esta medida someteríamos a algunos pacientes a biopsias prostáticas sin que tengan cáncer, falsos positivos, pero se nos escaparían menos cánceres; teniendo en cuenta la baja morbilidad de la biopsia prostática con respecto al carcinoma de prostata, esto sería deseable a no ser que se demostrara que el diagnóstico de estos cánceres no aumenta la supervivencia o que encontremos índices como densidad del PSA (PSAD) con mejor rendimiento diagnóstico en este intervalo de valores.

Así mismo la sensibilidad de la batería de test es el producto de las sensibilidades al aplicar cada uno de los test, mientras que el valor predictivo positivo de la batería será el obtenido al aplicar el último test de la misma.

La especificidad de la batería de test resultará del cociente de todos los casos denominados como negativos por cada uno de los test dividido por el número de casos de la enfermedad.

6.2. Test simultáneos

Otra posibilidad en la aplicación de test diagnósticos es cuando aplicamos más de un test, por ejemplo ecografía, tacto rectal y PSA, a un grupo de pacientes con y sin la enfermedad, carcinoma de próstata. En este caso un paciente se considerará como positivo cuando uno de los test aplicados lo encuadre como enfermo, con carcinoma de próstata. De esta manera disminuirémos los falsos negativos, por lo que aumentaremos la sensibilidad; lo que significa que la aplicación simultánea de test puede ser una estrategia adecuada para aplicarla como inicio de una secuencia. Es decir podríamos primero aplicar simultáneamente tacto rectal y PSA y cualquier paciente que tuviera tacto sospechoso o $PSA > 4 \text{ ng/ml}$ considerarlo como positivo y someterlo al siguiente test que podría ser la biopsia prostática; con esta estrategia obtendríamos menos falsos negativos que aplicando solo el PSA o el tacto rectal, pero la especificidad puede disminuir al aumentar los falsos positivos, aunque este hecho no es significativamente importante cuando no se aplican muchos test simultáneamente. A partir de la aplicación de trece test, la especificidad disminuye por debajo de 0.5.

C. INFERENCIA ESTADISTICA. EL ERROR ALEATORIO

E. Sánchez-Cantalejo Ramírez

1. Introducción

En el capítulo II.A se han presentado algunas herramientas que permiten describir los datos recogidos de una serie concreta de pacientes, permitiendo cuantificar tanto su magnitud como su variabilidad. Sin embargo, en el proceso de la investigación clínica los pacientes de interés no son sólo aquellos pacientes concretos sobre los que hemos medido las variables sino un grupo mucho más numeroso; así, a la hora de probar la posible eficacia de un tratamiento contra una enfermedad lo vamos a llevar a cabo sobre un grupo de 100 o 1000 pacientes, pero el interés está en poder saber si el tratamiento es eficaz en «todos» los pacientes con tal enfermedad. Así, es deseable disponer de herramientas que nos permitan generalizar los resultados encontrados en pequeños grupos de pacientes, como son los que se dispone habitualmente en la investigación clínica, a grupos de pacientes más numerosos. En este capítulo se presentan dos de tales herramientas, los intervalos de confianza y los tests o contrastes de hipótesis, que juegan un papel fundamental en la denominada Inferencia Estadística.

2. De lo particular a lo general

Cuando un clínico se plantea un estudio para evaluar uno o varios tratamientos, para establecer relaciones entre variables, etc., su interés no suele estar exclusivamente en los pacientes concretos a los que él ha tenido o puede tener acceso, si no más bien en todos los pacientes similares a estos. Al conjunto de pacientes en los que se quiere estudiar alguna cuestión se denomina, en términos estadísticos, la **población** y a los individuos concretos que participan en el estudio se denomina la **muestra**; al número de individuos en la muestra se conoce con el nombre de **tamaño de la muestra**. Por tanto, el clínico va a tener acceso a los individuos de la muestra y sin embargo en lo que está interesado es en estudiar el problema en la población. Imaginemos a un urólogo que quiere conocer el tiempo medio de la duración del tratamiento mediante litotricia, expresada en minutos; para ello tomará una muestra de, por ejemplo, 100

pacientes con litiasis, medirá en cada uno de ellos el tiempo que duró su tratamiento de litiasis y calculará el tiempo medio en esos 100 individuos; imaginemos que tal valor medio es 53.77 minutos; ¿qué decir acerca del tiempo medio en **todos** los individuos sometidos a litotricia?. El conjunto de procedimientos estadísticos que permiten pasar de lo particular, la muestra, a lo general, la población, es lo que se denomina la **Inferencia estadística**.

Es evidente que como de lo que se trata es de, a partir de lo observado en la muestra, tratar de saber lo que pasa en la población, se deben establecer unos criterios de elección de los individuos de la muestra, de tal forma que estos reflejen las características generales de la población, dicho en otros términos, que la muestra sea representativa de la población. Si tomamos como sujetos de la muestra sólo a hombres con más de 50 años, es posible que la media de estos sujetos no dé una buena idea del tiempo medio que dura el tratamiento con litotricia, es decir, la muestra no es representativa de la población. En estadística existen un conjunto de procedimientos, que se conocen con el nombre de **técnicas de muestreo**, que permiten elegir muestras de tal forma que éstas reflejen las características de la población; para llevar a cabo estas técnicas de elección de muestras es necesario en muchas ocasiones disponer de un listado de todos los individuos de la población. Por ejemplo, si se está interesado en evaluar el porcentaje de historias clínicas bien cumplimentadas de un servicio, podemos disponer de un listado de tales historias; si fuese muy costoso examinar cada una de esas historias, lo que se puede hacer es tomar una muestra de ellas elegidas por algún procedimiento de muestreo.

El hecho de no estudiar el problema de interés en toda la población sino sólo en una parte de ella, la muestra, conlleva un error; si la muestra se elige mediante algún proceso dependiente del azar, es decir, si se hace un **muestreo probabilístico o aleatorio**, se tiene la ventaja de que de alguna manera se puede evaluar el error antes mencionado. Si por el contrario, la muestra se elige a nuestra conveniencia, no hay forma de medir el error que se está cometiendo. El error asociado a estudiar un problema en base a una muestra elegida al azar se denomina **error aleatorio** para diferenciarlo de otros tipos de errores que se pueden cometer en un estudio y que se han descrito antes en esta monografía. Como quiera que casi nunca vamos a poder estudiar a toda la población, el error aleatorio es un error inevitable.

El clínico casi nunca va a disponer de un listado de los individuos de la población de su interés por lo que no va a poder hacer casi nunca ningún tipo de muestreo aleatorio; va a utilizar como muestra a los individuos disponibles, en general, sus pacientes o pacientes de colegas y lo que supondrá, para poder aplicar los métodos estadísticos inferenciales, es que sus pacientes forman una muestra aleatoria de una población que no se puede definir exactamente. A partir de lo que se observe en la

muestra, mediante algún proceso inferencial, tratará de conocer lo que ocurre en la población. La Fig. II.19 representa el proceso que se acaba de describir.

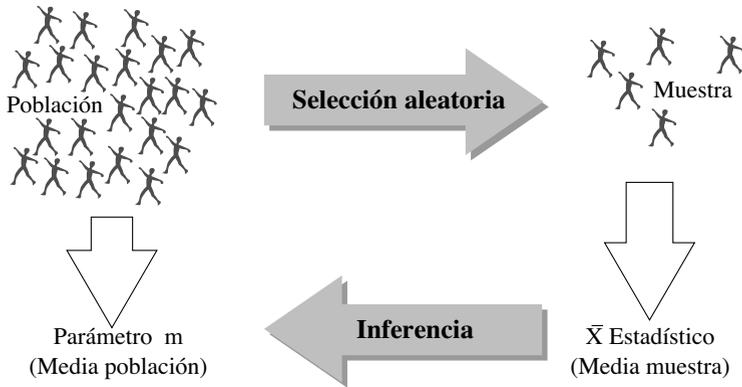


Fig. II.19.- El proceso de la inferencia estadística.

Como se ha dicho antes, la imposibilidad de estudiar a todos los individuos de la población de interés obliga al clínico a trabajar con muestras, lo que lleva inevitablemente a un error aleatorio. Veamos la repercusión que implica el hecho de trabajar con la información parcial que proporciona la muestra; siguiendo con el ejemplo anterior acerca de conocer el tiempo medio de duración del tratamiento con litotricia, antes se tomó una muestra de 100 pacientes, en los cuales el tiempo medio fué de 53.77 minutos; de este valor se dice que es una **estimación** del tiempo medio en la población de todos los pacientes tratados con litotricia, el parámetro a estimar; en general, un parámetro de la población es alguna característica, la media, varianza, etc. que defina la población en estudio. Si en lugar de haber tomado esta muestra de 100 pacientes hubiésemos tomado otra muestra aleatoria distinta, es decir, si repitiésemos la experiencia anterior, también con 100 pacientes, es fácil aceptar que el valor de la media en la nueva muestra no sería exactamente 53.77 minutos; supongamos que la media en la nueva muestra hubiese sido 57.08 minutos. Si realizamos esta experiencia muchas veces, es decir, tomamos muchas muestras de tamaño 100, tendremos tantas medias como muestras se hayan tomado.

Lo primero que salta a la vista después de esto es que no todas las medias de las muestras son iguales; la media de la muestra y, en general, cualquier forma de combinar los valores de la muestra para estimar un parámetro, lo que en estadística se denomina un **estimador**, cambia su valor de una muestra a otra. La Fig. II.20 es un histograma de la distribución de los valores de las medias de 200 muestras de tamaño 100, elegidas al azar de una población constituida por 3376 pacientes sometidos a tratamiento con litotricia; en esta población, la media es 57.46 minutos y la desviación típica 22.84 minutos. De la Fig. II.20 podemos observar que la mayoría de las medias

de las muestras están entre 54 y 60 minutos, es decir, alrededor de la media de la población que es 57.46 minutos; por otra parte, la media más pequeña entre las 200 calculadas es mayor que 52 minutos y la media más grande es menor que 65 minutos.

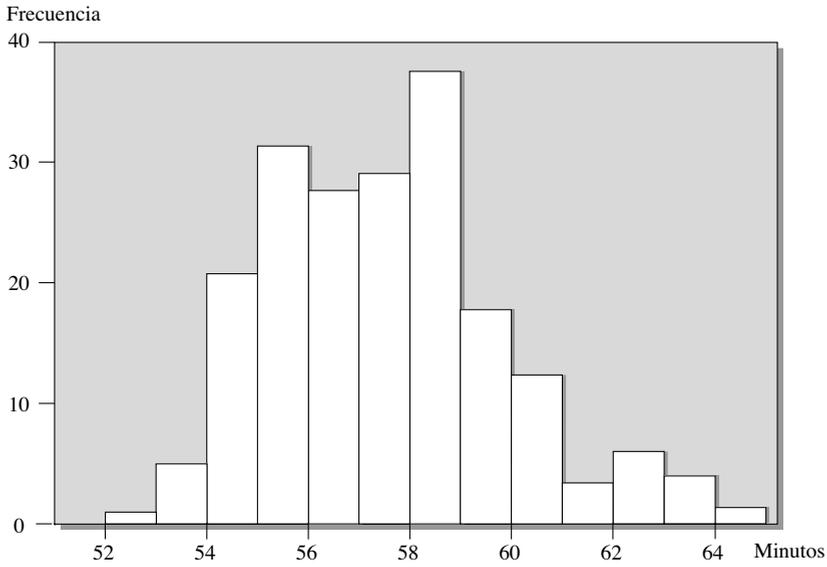


Fig. II. 20.- Distribución del tiempo medio del tratamiento con litotricia en base a 200 muestras de 100 enfermos.

3. Valor de la información de nuestra experiencia

Ya que ahora somos conscientes de que el resultado de nuestra experiencia no es más que uno de los resultados que se podrían haber obtenido, la pregunta a responder es: ¿cómo de fiable es el valor de la estimación de nuestra experiencia concreta?. El **error estándar de la media** es la medida que trata de captar la variabilidad de los valores de la media en todas las posibles muestras de un tamaño determinado, en nuestro caso 100; en general, el error estándar de cualquier estimador, una media, una proporción, etc., mide el grado de variabilidad en los valores del estimador en las distintas muestras de un determinado tamaño que pudiésemos tomar de una población. Cuanto más pequeño sea el error estándar de un estimador más nos podremos fiar del valor de una muestra concreta; si en lugar de variar las 200 medias entre los valores de 52 a 65 minutos hubiesen variado entre 20 y 90 minutos sería menos probable que, al tomar una muestra concreta y calcular su media, ésta estuviese cercana al valor 57.46 que es el valor de la media de la población, en nuestro caso, el parámetro de interés. La expresión del error estándar depende del estimador de que se trate, pero su cálculo es lo de menos pues nos lo ofrece cualquier paquete estadístico que utilicemos al

analizar los datos; para el caso de estimar la media de la población, una estimación del error estándar de la media viene dada por

$$s/\sqrt{n}$$

siendo s la desviación típica de la muestra y n el tamaño de la muestra.

Si en lugar de tratar de estimar la duración media del tratamiento con litotricia estuviésemos interesados en conocer la proporción de hombres con cáncer de próstata que tienen afectación ganglionar, obtendríamos muestras aleatorias de la población y en cada una de ellas podríamos calcular la proporción de hombres con afectación ganglionar, estimación de la proporción en la población. Pues bien, si representásemos cada una de las proporciones muestrales se repetiría un hecho similar al observado para el caso de las medias: la mayoría de las proporciones en las distintas muestras estarían próximas a la proporción en la población y la representación gráfica mediante un histograma de las proporciones muestrales tendría una forma semejante. Para el caso de una proporción, una estimación del error estándar es

$$\sqrt{p(1-p)/n}$$

donde p es la proporción en la muestra. Sin embargo, hay situaciones en que el estimador de un parámetro no se distribuye de la misma forma que para el caso de la media o de la proporción; ejemplos de ello son la razón del producto cruzado (O.R.) y el riesgo relativo. Estos dos estadísticos no tienen una distribución simétrica como ocurre para el caso de la media y de la proporción; sin embargo, si se toma su transformación logarítmica su distribución se hace simétrica y además se conoce la expresión de los errores estándar; así, el error estándar para el logaritmo del riesgo relativo viene dado por la expresión

$$\sqrt{\frac{b}{a(a+b)} + \frac{d}{c(c+d)}}$$

donde a, b, c, d son los valores de las casillas, como aparecen en la Tabla I.1. Así, para los resultados del estudio sobre tumor de Wilms que aparecen en la Tabla I.2, el error estándar para el logaritmo del riesgo relativo será

$$\sqrt{\frac{21}{2(2+21)} + \frac{234}{2(2+234)}} = 0.976$$

De forma similar, el error estándar para el logaritmo de la O.R. viene dado por la expresión

$$\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

Obsérvese como en el error estándar de cualquier estimador, el tamaño de la muestra figura en el denominador, lo que significa que cuanto mayor sea el tamaño de la muestra menor será el error estándar; esto es un hecho que ocurre para todos los estimadores: a mayor tamaño de muestra, mayor precisión.

Esto que se puede probar empíricamente es un resultado muy importante de la estadística que se conoce como el **teorema central del límite** que se puede enunciar así: para los estimadores que pueden ser expresados como suma de los valores muestrales, la distribución de sus valores en las distintas muestras sigue aproximadamente una distribución Normal con media la de la población y desviación típica igual al error estándar del estimador del que se trate. La distribución Normal depende de dos parámetros, la media y la desviación típica y tiene la forma de «campana»; es una distribución muy importante en estadística no porque muchas variables biomédicas sigan tal distribución sino por el papel primordial que juega el teorema central del límite en toda la estadística; no hace falta aclarar que la palabra «Normal» para esta distribución nada tiene que ver con el sentido de normalidad biológica. El 90% de los valores de una distribución Normal se encuentran en el rango: media \pm (1.64) (desviación típica); el 95% de los valores están en el rango «media \pm (1.95) (desviación típica)» y el 99% de los valores en el rango «media \pm (2.58) (desviación típica)». En la Fig. II.21 aparece la distribución de las medias de las muestras que se obtendría si se hubiesen elegido todas las posibles muestras de tamaño 100.

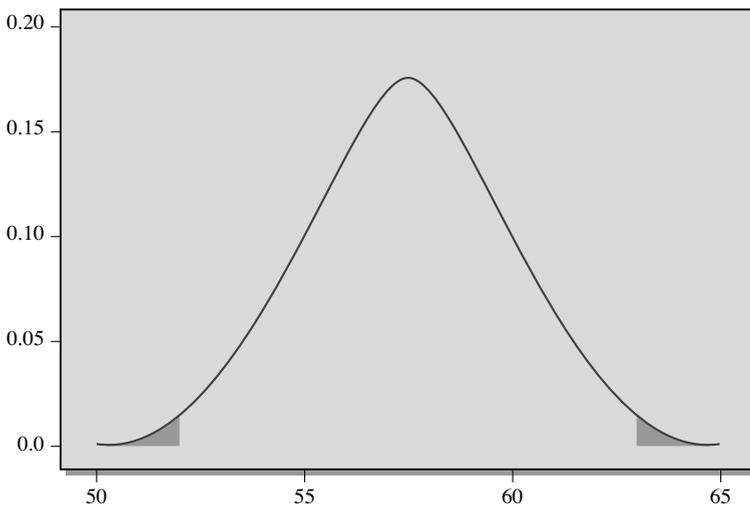


Fig. II. 21.- Distribución de las medias de todas las posibles muestras de tamaño 100.

4. Control del error aleatorio: intervalo de confianza

El ejemplo que acabamos de exponer nos ha permitido introducir algunos conceptos fundamentales y cada vez más frecuentes en la literatura clínica, pero es irreal en el sentido de que nunca vamos a conocer el parámetro de interés en la población. Entonces cuando el urólogo estudie sus individuos concretos y calcule una media de la duración, no va a saber si la media de su muestra está o no cercana a la media de la población, es decir, no va a saber si la media de todos los pacientes similares a estos van a tener una media parecida.

Los **intervalos de confianza** son un medio de conocer el valor de un parámetro en una población, midiendo el grado de error implícito en todo proceso que tiene que ver con el azar. Un intervalo de confianza para un parámetro de una población no es más que una pareja de números A, B tales que, con un nivel de confianza determinado, podemos asegurar que el valor del parámetro es mayor que A y menor que B. En general, si el estimador muestral sigue una distribución Normal, un intervalo de confianza para un parámetro de una población viene dado por la expresión

$$\text{estimador} \pm z (e.e.)$$

donde *e.e.* es el error estándar del estimador y z es un valor que depende del **nivel de confianza** $1-\alpha$ con que se quiera dar el intervalo; los valores de z para los niveles de confianza 0.90, 0.95 y 0.99 son 1.64, 1.96 y 2.58 respectivamente; el signo \pm significa que cuando se elija el signo negativo se conseguirá el extremo inferior del intervalo, y cuando se elija el signo positivo se tendrá el extremo superior. A la diferencia entre 1 y el nivel de confianza se le denomina **nivel de error**.

Pero, ¿qué significa que la confianza de un intervalo es del 95%? Para ello vamos a volver al caso de querer conocer el tiempo medio de tratamiento mediante litotricia; a partir de cada muestra elegida podemos calcular el valor del estimador, la media de la muestra, y además, podemos estimar, hacernos idea de, el error estándar de la media; según la expresión anterior, es todo lo que necesitamos, además de sustituir z_α por 1.96, para calcular el intervalo de confianza con objeto de conocer el valor medio en la población. Pues bien, si tomamos muchas muestras, podemos calcular otros tantos intervalos; por un argumento estadístico que cae fuera del nivel de esta monografía, se puede demostrar que por término medio, de cada 100 de tales intervalos, 95 de ellos contendrán en su interior el valor del parámetro en la población; en la Fig. II.22 aparecen 100 intervalos procedentes de las primeras 100 muestras de las 200 que se tomaron y como se observa la mayoría, 92 de los 100 los intervalos, contienen en su interior el valor 57.46 que era la media en la población; si esto lo hiciésemos muchas veces, de cada 100 intervalos, 95 de ellos, por término medio, contendrán en su interior el valor 57.46 del parámetro.

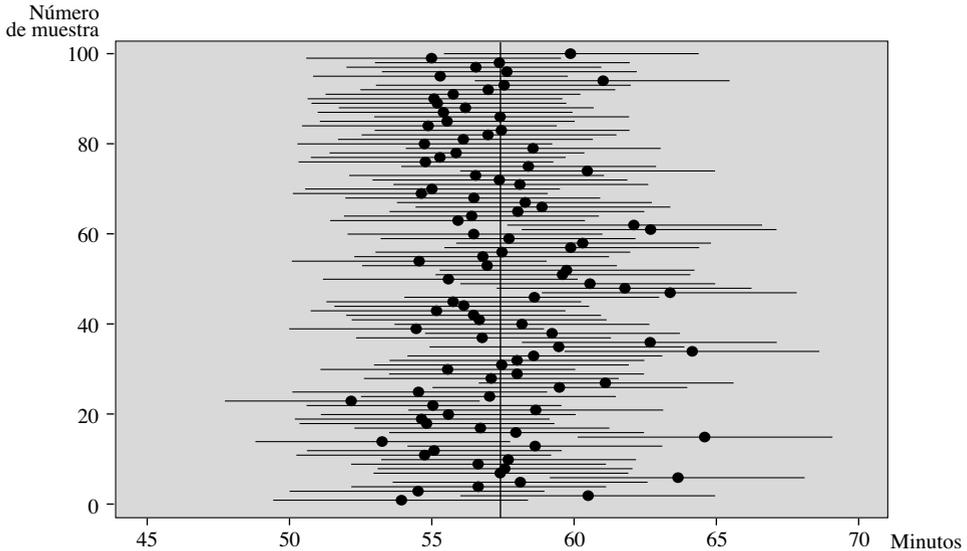


Fig. II. 22.- Representación de 100 intervalos, al 95% de confianza para la duración media de tratamiento con litotricia.

Evidentemente, en la práctica no nos vamos a dedicar a tomar muchas muestras, tan sólo tomaremos una; a partir de ella calcularemos la media y el error estándar y obtendremos el intervalo a la confianza elegida; supongamos que hemos tomado una muestra que dió una media de 58.17 minutos y un error estándar de 3.09 minutos; de aquí, un intervalo, al 95% de confianza para la media en la población será

$$58.17 \pm 1.96 (3.09) = (52.11 , 64.23)$$

¿Cómo interpretar este resultado? Esperamos que éste sea uno de los 95 intervalos de cada 100 que incluyen en su interior el valor de la media de la población; sería mala suerte que este fuese uno de los 5 de cada 100 que no tienen en su interior el valor medio de la población; de esta interpretación debe quedar claro que tal resultado no significa que existe una probabilidad del 95% de que la media de la población esté en su interior. En definitiva, la información que presta un intervalo de confianza para un parámetro es un par de números entre los que nosotros confiamos que esté el valor del parámetro de interés; de ninguna manera lo podemos asegurar con certeza total, pero estamos casi seguros.

Este grado de confianza o seguridad lo podemos elegir nosotros por lo que el lector puede pensar que porqué no construir el intervalo a un mayor nivel de confianza; el intervalo anterior al 99% de confianza resulta de sustituir el 1.96 que corresponde al 95%, por el 2.58 que corresponde al 99%; por tanto, el intervalo es ahora

$$58.17 \pm 2.58 (3.09) = (50.20 , 66.14)$$

intervalo más ancho que el anterior pues si queremos tener más confianza en que nuestro intervalo contenga a la media de la población es evidente que debe ser más ancho. Los valores más comunes de niveles de confianza son los antes mencionados, 0.90, 0.95 y 0.99.

En el apartado anterior se calculó el error estándar para el logaritmo del riesgo relativo para el estudio cuyos datos aparecen en la Tabla I.2; por tanto, ya que una estimación del riesgo relativo fue de 10, un intervalo, al 95% de confianza, para el logaritmo del riesgo relativo será

$$\log (10) \pm 1.96 (0.976) = (0.390 , 4.216)$$

donde $\log (10)$ es el logaritmo neperiano de 10; por tanto, el intervalo para el riesgo relativo será

$$(e^{0.39} , e^{4.216}) = (1.48 , 67.76)$$

Parece claro que cuanto más estrecho sea el intervalo más precisa es la información que proporciona acerca del parámetro de interés; por tanto la pregunta a responder es: ¿cómo hacer el intervalo lo más estrecho posible?. Una forma sería disminuyendo la confianza, a riesgo de aumentar las posibilidades de error; una segunda alternativa para estrechar el intervalo es disminuir el error estándar del estimador y esto se puede conseguir aumentando el tamaño de la muestra elegida, pues cuanto mayor sea el tamaño de las muestras, más parecidas serán las distintas estimaciones a partir de ellas, es decir, menor será el error estándar del estimador del que se trate. El problema del cálculo del tamaño de muestra a elegir para estimar un parámetro mediante un intervalo de confianza, es en ocasiones algo complejo por lo que el clínico debería consultar con un estadístico.

Esta discusión que se ha hecho con el intervalo de confianza para la media de una población, se puede hacer para cualquier otro parámetro: una prevalencia o proporción de sujetos de una población que tienen una determinada característica, diferencias de medias o de proporciones, una O.R., un riesgo relativo, etc.; lo único que cambia es que las expresiones del intervalo son algo más complicadas. De todas formas, como se dijo antes, el ordenador nos evita tener que conocer tales fórmulas pues él ya nos da los cálculos directamente.

Imaginemos un estudio de casos y controles para evaluar si existe o no relación entre la presencia de retención aguda de orina en pacientes con adenoma de próstata y la puntuación obtenida en un I-PSS (cuestionario de síntomas propuesto por la

O.M.S.); si hemos obtenido una estimación de la razón de producto cruzado (*O.R.*) de 1.56 con un intervalo, al 0.95 de confianza, comprendido entre 1.11 y 1.78; esto quiere decir que estamos casi seguros, tenemos una confianza del 95%, que el intervalo anterior contenga en su interior a la *O.R.* de la población de todos los posibles pacientes similares a estos; en otras palabras, con una confianza del 95%, podemos afirmar que la *O.R.* de la población es como mínimo 1.11 y como máximo 1.78. Si el intervalo hubiese ido de 0.77 a 2.11, ello significa que el valor 1 es un valor plausible para la *O.R.* de la población y ya que una *O.R.* que vale la unidad significa la no asociación, es decir, la independencia, entre las dos variables en juego, nos llevaría a establecer que no hay evidencia para afirmar que las dos variables estén asociadas.

Resumiendo, un intervalo de confianza da un rango de valores posibles para el parámetro poblacional. En la literatura médica es mucho más frecuente el uso de los contrastes de hipótesis como herramienta inferencial que los intervalos de confianza; ya que estos dan información suplementaria, muchos autores, entre ellos Gadner y Altman¹⁴ recomiendan su uso en la investigación clínica.

5. Control del error aleatorio: los contrastes de hipótesis

Además de los intervalos de confianza, otra herramienta fundamental en el proceso de inferencia estadística son los llamados tests o contrastes de hipótesis. Con los intervalos podemos hacernos idea de un parámetro en una población dando un par de números entre los que confiamos que esté el valor del parámetro desconocido; los contrastes de hipótesis utilizan otra estrategia: se trata de establecer a priori una hipótesis acerca del valor de un parámetro, realizar un estudio y analizar la coherencia entre la hipótesis previamente establecida y los hallazgos encontrados en el estudio. Con objeto de facilitar la comprensión de este proceso argumental, vamos a presentar un contraste de hipótesis, ejemplificando con datos reales.

5. 1. Un ejemplo introductorio

Mora et al.¹⁵ presentan un estudio sobre el efecto del Flurbiprofen frente a la combinación de Dipirona e Hioscina para el tratamiento del dolor en el cólico nefrítico. Para ello diseñaron un ensayo clínico en el que hay dos grupos de enfermos de dolor crónico, uno que recibe Flurbiprofen mientras a los enfermos de la segunda cohorte se les administra Dipirona e Hioscina. Como medida del efecto de los tratamientos se definió la respuesta, positiva o negativa, a un tratamiento cuando se consiguió o no una situación de alivio. Los resultados del estudio aparecen en la Tabla II.11. Este es el ejemplo más simple de las llamadas **tablas de contigencia o tablas cruzadas**, donde aparecen en las filas de la tabla las distintas categorías de una variable categórica, en

este caso el tipo de tratamiento, y por columnas las categorías de otra variable, en este caso el tipo de respuesta al dolor producido por un cólico nefrítico. A la Tabla II.11 se le denomina **2x2** porque cada una de las dos variables tiene solo dos categorías.

Tabla II.11.- Respuesta al dolor en 52 enfermos con cólico nefrítico.

	Respuesta al dolor		Total
	Positiva	Negativa	
Dipirona +Hioscina	11	15	26
Flurbiprofen	16	10	26
Total	27	25	52

Según estos resultados, de los 26 enfermos que recibieron Flurbiprofen, 16 presentaron respuesta positiva, es decir, el 61.5% respondieron bien a este medicamento, mientras que sólo el 42.3% lo hicieron a la combinación Dipirona+Hioscina; está claro que en este grupo de 52 pacientes, el Flurbiprofen ha sido más efectivo que el otro tratamiento. Sin embargo, cuando los investigadores se plantearon el estudio en lo que estaban interesados era en la efectividad del Flurbiprofen no en estos 52 individuos concretos sino en el conjunto de enfermos con cólico nefrítico con condiciones similares a éstos. Por tanto se plantea la siguiente pregunta: ¿qué hubiese ocurrido si en vez de tomar estos 52 hubiésemos tomado otros 52 enfermos distintos?, ¿se mantendrían esas diferencias de efectividad a favor del Flurbiprofen o cambiarían?; porque es fácil admitir que aunque los dos tratamientos que se comparan son igualmente efectivos, es muy probable que no hubiesen tenido exactamente la misma proporción de respuestas positivas. Puestas así las cosas, ¿es posible achacar al azar la diferencia encontrada entre las proporciones de respuestas positivas? O si, por el contrario, en esa diferencia hay algo más que azar y si el estudio estuvo bien diseñado, ¿habrá que concluir que los dos tratamientos no son igualmente efectivos?.

5.2. Tests o contrastes de hipótesis

Los **tests o contrastes de hipótesis** son herramientas estadísticas que permiten responder a preguntas como la que se acaba de plantear; un contraste de hipótesis permite cuantificar la compatibilidad entre una hipótesis previamente establecida con los resultados obtenidos en un estudio. Si como resultado del contraste resultara que existe incompatibilidad entre la hipótesis previa y los datos del estudio, la decisión más razonable sería rechazar lo que se estableció previamente; en caso contrario, no debe haber inconveniente en aceptarla.

Sean cuales sean los deseos de los investigadores del estudio anterior, en todo contraste de hipótesis se parte de la que se denomina **hipótesis nula** que, para un estudio como el anterior, establece la igualdad en la efectividad de los dos tratamientos; si p_{D+H} representa la probabilidad de tener respuesta positiva al tratamiento con Dipirona + Hioscina y p_F la correspondiente al Flurbiprofen, la hipótesis nula a contrastar establece que $p_{D+H} = p_F$. Por tanto, lo que vamos a hacer es medir la compatibilidad de esta igualdad con el hecho de haber encontrado un 42.3% y 61.5% de respuestas positivas en los grupos.

El razonamiento a seguir es el siguiente: suponer la hipótesis nula cierta y estudiar cómo es de probable que siendo iguales los dos tratamientos se obtengan resultados como los obtenidos o haber encontrado diferencias más grandes entre los grupos. Siguiendo con el ejemplo anterior, si los dos tratamientos fuesen iguales, una estimación de la probabilidad de tener respuesta positiva al dolor sería $27/52$; por tanto, supuesto que los dos tratamientos fuesen iguales, el número esperable de enfermos con respuesta positiva de entre los que recibieron Dipirona+Hioscina sería $(27/52)(26)=13.5$. En definitiva, en el primer grupo de enfermos hemos observado 11 respuestas positivas al dolor y acabamos de ver que si los dos tratamientos fuesen iguales, hubiésemos esperado 13.5. Este valor 13.5 se ha calculado de la siguiente forma: $(27/52)(26)$, lo que quiere decir que lo esperado en la casilla (1,1), pues está en la primera fila y columna de la tabla, si la hipótesis nula fuese cierta, se puede conseguir multiplicando el total de la primera fila, 26, por el total de la primera columna, 27, y dividiendo por el total de individuos, 52.

De la misma forma, para cada casilla de la tabla, podemos calcular el número esperable de enfermos, bajo la hipótesis de que los dos tratamientos fuesen iguales; por tanto, podemos formar una nueva tabla, la llamada **tabla de valores esperados**, que refleja lo que debería haber ocurrido si los dos tratamientos fuesen iguales.

Tabla II.12.-Valores esperados en caso de igualdad de los dos tratamientos.

	Respuesta al dolor		Total
	Positiva	Negativa	
Dipirona +Hioscina	13.5	12.5	26
Flurbiprofen	13.5	12.5	26
Total	27	25	52

El hecho de que los dos tratamientos tengan iguales valores esperados se debe exclusivamente a que los dos grupos de tratamiento tienen el mismo número de enfermos.

En definitiva, lo que se acaba de hacer es construir una tabla que es una traducción de la hipótesis nula, la que establece la igualdad en la efectividad de los dos tratamientos. Como quiera que el objetivo del contraste de hipótesis era medir la compatibilidad entre la tabla de los valores observados y la hipótesis nula, tal objetivo se puede conseguir midiendo la compatibilidad entre la tabla de los valores observados y la de los valores esperados.

La pregunta que ahora se plantea es ¿cómo medir la compatibilidad entre esas dos tablas?. El estadístico británico K. Pearson propuso la cantidad

$$\frac{(11-13.5)^2}{13.5} + \frac{(15-12.5)^2}{12.5} + \frac{(16-13.5)^2}{13.5} + \frac{(10-12.5)^2}{12.5} = 1.93$$

como medida de la similitud entre la tabla de valores observados y la de valores esperados. Obsérvese que lo que se hace es comparar cada casilla de la primera tabla con la de la segunda; por tanto si las dos tablas fuesen iguales, cada uno de los numeradores de las anteriores fracciones sería cero por lo que la cantidad anterior sería cero. Por otra parte, si las dos tablas fuesen muy diferentes, los numeradores serían grandes y la suma de las fracciones también sería grande.

Ya que antes se ha dicho que la tabla de valores esperados es un reflejo de la hipótesis nula, cuando la cantidad anterior sea grande y por tanto nos indique incompatibilidad entre las dos tablas, realmente nos estará hablando de la incompatibilidad entre la tabla de valores observados y la hipótesis de igualdad de los dos tratamientos, la hipótesis nula. Si fueran así las cosas, al existir incompatibilidad entre la hipótesis establecida y los valores observados a partir del estudio realizado, la solución más racional sería el rechazo de la hipótesis nula y por tanto la conclusión de que los dos tratamientos no son igualmente efectivos. Si, por el contrario, la cantidad propuesta por Pearson fuese pequeña, nos hablaría de la similitud entre las dos tablas, o lo que es igual, la compatibilidad entre lo observado y la hipótesis nula; ya que la hipótesis de igualdad de los dos tratamientos no entra en contradicción con lo observado, no hay motivos para rechazar tal hipótesis, por lo que concluiríamos aceptándola.

En definitiva, el estadístico de Pearson es una medida que nos permite decidir si rechazar o no la hipótesis de igualdad de los dos tratamientos; en caso de que sea pequeña no tenemos argumentos para rechazar la hipótesis nula, o lo que es igual, aceptare-

mos la igualdad de los tratamientos y en caso contrario diremos que son distintos. Lo que queda por definir es el criterio para decidir si el estadístico de Pearson es pequeño o grande; tal decisión la vamos a tomar en función del denominado **valor P**. Este valor nos da la probabilidad de que, supuesto que los dos tratamientos son igualmente efectivos, encontremos en estudios similares al nuestro resultados como los que hemos observado o resultados más contradictorios con la hipótesis nula. Si esta probabilidad es pequeña, en general, menor de 0.05 ($P < 0.05$), querrá decir que es muy poco probable que siendo la hipótesis nula verdadera, se encuentren resultados como los observados o más extremos; dicho de otra forma, lo observado es incompatible con la hipótesis nula y por tanto la rechazamos y concluimos que los dos tratamientos no son igualmente efectivos. Cuando se rechaza la hipótesis nula se suele decir que **el test es significativo**. Si, por el contrario, el valor P es grande significa que es posible que siendo los dos tratamientos igualmente efectivos, los hallazgos de nuestro estudio no son raros, por lo que no habrá razones para rechazar la hipótesis nula y por tanto la aceptaremos. J. H. Ware et. al.¹⁶ definen el valor P como el grado de sorpresa; cuanto más pequeño sea P, más sorprendentes son nuestros resultados en el caso en que la hipótesis nula sea cierta.

Como cualquier test estadístico, solo se puede aplicar en ciertas condiciones; para el caso del test de la χ^2 que se acaba de exponer es aplicable siempre que el número de sujetos en el estudio sea al menos 40; si éste es menor que 40 pero mayor que 20, ningún valor esperado debe ser menor que 5. Si no se cumple esto, se puede aplicar el llamado test de Fisher¹⁷.

5. 3. Tipos de error

Es posible que, aún siendo verdadera la hipótesis nula, la diferencia entre las proporciones de respuestas positivas observadas entre los dos tratamientos sea grande debido al azar; es posible, por ejemplo, que al lanzar dos monedas perfectas 10 veces al aire con una de ellas obtengamos 1 cara y 9 con la otra, aunque esto es ciertamente improbable; en resumen, existe la posibilidad de que siendo verdadera la hipótesis nula, el azar nos lleve a rechazarla. Este error, consistente en rechazar la hipótesis nula siendo verdadera se conoce como **error de tipo I** o **error α** . Por otra parte, es posible que los resultados experimentales nos lleven a aceptar la hipótesis nula siendo falsa; a este nuevo error se le conoce como **error de tipo II** o **error β** . Si representamos por H_0 a la hipótesis nula y por H_1 a la hipótesis contraria, la que establece que los dos tratamientos tienen distinta efectividad, el siguiente cuadro presenta las cuatro posibles situaciones en las que nos podemos encontrar.

Tabla II. 13. Los cuatro posibles resultados de un contraste de hipótesis.

		Decisión por	
		H_0	H_1
Hipótesis cierta	H_0	Decisión correcta	Error tipo I (α)
	H_1	Error tipo II (β)	Decisión correcta

Como quiera que no conocemos cuál de las dos hipótesis es la verdadera, aceptemos la hipótesis nula o la rechazemos, siempre corremos el riesgo de equivocarnos, sea cual sea la decisión que tomemos.

A la cantidad $1-\beta$, la probabilidad de rechazar la hipótesis nula siendo falsa, se denomina **potencia** del test; dicho de otra forma, la potencia del contraste antes empleado se puede definir como la verosimilitud de que, supuesto de que los dos tratamientos son distintos, el test lo detecte. Si el contraste realizado acerca del Flurbiprofen y la combinación de Dipirona+Hioscina tuviese una potencia del 90%, ello significaría que existe una probabilidad del 90% de detectar diferencias entre los tratamientos, supuesto que tal diferencia exista. Cuando un estudio da lugar a resultados no significativos es muy importante calcular la potencia del test empleado en detectarla; como luego se verá, cuando el tamaño de muestra decrece, también lo hace la potencia del contraste.

Entre los errores de tipo I y II hay una íntima relación; para un mismo tamaño de muestra, cuando se disminuye el error α , es decir, deseamos disminuir la probabilidad de equivocarnos al rechazar la hipótesis nula, el error β , la probabilidad de aceptar la hipótesis nula siendo falsa, aumenta. Debe quedar muy claro por otra parte que, *con una misma muestra*, puede aceptarse o rechazarse la hipótesis nula; todo depende del error α considerado de tal forma que cuanto más pequeño sea más difícil será rechazarla pero más probable su aceptación siendo falsa. Pues bien, el error α más pequeño al que se puede rechazar la hipótesis nula es lo que antes hemos denominado error P; supongamos que una hipótesis se puede rechazar para un nivel de error α del 5% pero no para un $\alpha=1\%$; esto significa que el nivel de error más pequeño al que se puede rechazar H_0 debe ser un número comprendido entre 0.05 y 0.01 y esto se suele escribir de la forma

$$P < 0.05$$

Esta notación proviene del hecho de que antes de la difusión de los paquetes estadísticos-informáticos, el estadístico solo tenía acceso a tablas en las que únicamente

aparecían determinados valores de α ; los más usuales son: 0.20, 0.10, 0.05, 0.01 y 0.001. Por esta razón no se podía calcular el valor exacto de P y había que conformarse con decir que era menor que 0.05 y mayor que 0.01; sin embargo, con la popularización de los ordenadores el investigador puede obtener, para cualquier test, el error P exacto por lo que la notación anterior irá desapareciendo con el tiempo. Para el ejemplo que venimos siguiendo, el error mínimo al cual podemos rechazar la hipótesis de igualdad de los dos tratamientos es $P=0.16$; ya que este es un error bastante grande respecto al convencional, el 0.05, no podemos rechazar la hipótesis nula y por tanto la aceptamos. En definitiva, a la vista de los resultados de este estudio no hay argumentos para decir que los dos tratamientos tienen distinta efectividad, lo que, por abuso del lenguaje, se suele querer decir con la frase: *los dos tratamientos tienen igual efectividad*. Aunque para cada caso particular habría que discutir el error α a elegir, la costumbre es rechazar la hipótesis nula si P es inferior a 0.05, lo que no se debe tomar como dogma.

Para estimar la diferencia de efectividad de los dos tratamientos lo podemos realizar mediante un intervalo para la diferencia de proporciones de éxitos en las dos poblaciones, los que recibieron Flurbiprofen y los que tomaron Dipirona+Hioscina; tal intervalo viene dado por la expresión

$$p_F - p_{D+H} \pm z_\alpha \sqrt{\frac{p_F(1 - p_F)}{n_F} + \frac{p_{D+H}(1 - p_{D+H})}{n_{D+H}}}$$

donde p_F y n_F son el porcentaje de éxitos y el tamaño de la muestra del Flurbiprofen y p_{D+H} y n_{D+H} la correspondiente a la Dipirona+Hioscina; sustituyendo lo encontrado en el estudio, un intervalo de confianza al 95% es

$$0.615 - 0.423 \pm 1.96 \sqrt{\frac{0.615(1 - 0.615)}{26} + \frac{0.423(1 - 0.423)}{26}} = (-0.075, 0.459)$$

Como se puede apreciar, este intervalo contiene en su interior el valor cero, lo que significa que éste puede ser un valor para la diferencia entre los dos tratamientos, es decir, los dos tratamientos pueden ser iguales, lo que está de acuerdo con los resultados del contraste de hipótesis que antes hemos realizado.

5. 4. Una posible limitación muy importante a los resultados anteriores

Cuando se comparan dos o más grupos de individuos para, por ejemplo, comparar distintos tratamientos, hay que ser muy cuidadosos a la hora de sacar conclusio-

nes acerca de la efectividad de los tratamientos . En el ejemplo que venimos mostrando en este capítulo, lo que realmente se ha probado es que es más probable que el Flurbiprofen cure a los enfermos de la hipotética población de la que provienen los enfermos que lo tomaron, que lo haga la Dipirona+Hioscina en la correspondiente población de la que proviene la muestra de enfermos que tomó tal combinación.

Si esas dos poblaciones no fueran iguales en términos de cualquier variable que pudiera afectar a la efectividad de los tratamientos, ¿a qué achacar la diferencia de efectividad de curaciones; a la diferencia entre los pacientes de los dos grupos o a los diferentes tratamientos?. Bien podría ocurrir que en el grupo del Flurbiprofen estuviesen los sujetos más jóvenes, con dolor menos intenso, etc. características que están a favor de una mayor probabilidad de curación.

Para responder a esta pregunta existen dos posibilidades. La primera consiste en diseñar un estudio, un ensayo clínico, de tal manera que los dos grupos fuesen parecidos en cuanto a variables que puedan afectar a la probabilidad de curación; la otra posible solución es medir la efectividad de los tratamientos teniendo en cuenta las posibles desigualdades entre los grupos mediante un análisis multivariante, cuyos principios básicos se presentan en el Capítulo II.E.

6. Tamaño de la muestra

Otra cuestión de vital importancia es el hecho de que, *para un error α fijo*, toda hipótesis nula se puede aceptar o rechazar dependiendo del tamaño de muestra elegido; con tamaños de muestra pequeños será difícil rechazar una hipótesis nula y con tamaños de muestra grandes será fácil rechazarla. Si estamos interesados en *demonstrar científicamente* que la eficacia de las dos drogas A y B es la misma, no tenemos más que aplicar ambos tratamientos a un número pequeño de individuos; lo que no se dice en estas ocasiones es que la potencia del test es muy pequeña, o lo que es igual, la probabilidad de aceptar H_0 cuando es falsa es muy grande. De esta forma parece claro que ante todo test no significativo, el caso de nuestro ejemplo, habría que preguntar por la potencia dando así al lector un elemento de juicio acerca de la verosimilitud de la hipótesis nula. Freiman et al.¹⁸, revisaron 71 ensayos clínicos con resultados no significativos publicados en *The New England Journal of Medicine* entre los años 1960 y 1977 y demostraron que la mitad de ellos tenían potencias menores de 0.26 (sus errores β eran mayores que 0.74) a la hora de detectar diferencias del orden del 25%. De esta forma se puede decir que si no se considera a la hora del diseño el tamaño de muestra, *la ausencia de una prueba de eficacia no es prueba de la ausencia de eficacia*.

Por el contrario, toda hipótesis nula, para un α dado, se puede rechazar sin más que tomar un número suficientemente grande de individuos en la(s) muestra(s); supongamos que se quiere realizar un test de hipótesis acerca de la eficacia de las dos drogas A y B; diferencias tan pequeñas como $p_{D+H} - p_F = 0.001$ pueden llevar el rechazo de $p_{D+H} - p_F = 0$ si tales estimaciones están basadas en grandes tamaños de muestra. Aquí aparece clara la necesidad de distinguir entre la **significación estadística** y la **significación práctica**; ¿hasta qué punto tiene sentido afirmar que dos drogas tienen efectos diferentes si sus probabilidades de curación respectivas son 0.5 y 0.501?; en el cálculo del tamaño de muestra interviene la diferencia que al investigador le parece importante detectar, es decir, la diferencia que tiene efectos prácticos.

En general, el tamaño de muestra a tomar depende de los errores α , β , de la mínima diferencia que se considera importante y de la variabilidad de la variable en cuestión. Veamos un ejemplo del cálculo del tamaño mínimo de muestra para el caso de la comparación de dos tratamientos cuya efectividad se va a medir mediante el porcentaje de éxitos; para calcular el tamaño de muestra deberemos conocer el porcentaje p_1 de éxitos del primer tratamiento, el tratamiento estándar, fijar la diferencia que debe tener el otro tratamiento respecto del estándar para que tal diferencia sea clínicamente relevante y elegir el nivel de error α y la potencia del contraste; definidas tales cantidades el tamaño mínimo de muestra viene dado por la expresión

$$n = \left(\frac{z_{\alpha} \sqrt{2p(1-p)} + z_{2\beta} \sqrt{p_1(1-p) + p_2(1-p_2)}}{p_1 - p_2} \right)^2$$

donde $p_1 - p_2$ es la diferencia que se tomó como clínicamente importante y $p = (p_1 + p_2)/2$; z_{α} y $z_{2\beta}$ son cantidades que dependen de los errores antes fijados.

Supongamos que la experiencia dice que el 40% de los tratados con la combinación Dipirona+Hioscina responde positivamente; vamos a establecer como diferencia clínicamente relevante la cantidad 0.1, es decir, consideraremos los dos tratamientos clínicamente distintos cuando se diferencien en al menos un 10% en el porcentaje de respuestas positivas; por tanto, $p_1 = 0.4$, $|p_1 - p_2| = 0.1$, es decir $p_2 = 0.3$ o $p_2 = 0.5$; por razones en las que no podemos entrar aquí siempre se debe tomar el valor más próximo a 0.5. Si elegimos como error $\alpha = 0.05$ y potencia $1 - \beta = 0.9$, es decir, $\beta = 0.1$, entonces $z_{0.05} = 1.96$ y $z_{2(0.1)} = z_{0.2} = 1.28$ según la tabla de la distribución normal, por lo que el tamaño de muestra viene dado por

$$n = \left(\frac{z_{\alpha} \sqrt{2p(1-p)} + z_{2\beta} \sqrt{p_1(1-p) + p_2(1-p_2)}}{p_1 - p_2} \right)^2$$

lo que quiere decir que debemos tomar al menos 518 pacientes por grupo. Si disminuimos la potencia a 0.8, el valor $z_{2\beta}$ es, según la tabla de la normal, 0.842 por lo que el tamaño de muestra necesario pasa a ser 388; este resultado de un menor tamaño de muestra era esperable pues lo que hemos hecho ha sido disminuir la potencia y, por tanto, aumentar el error de aceptar que los dos tratamientos son iguales cuando son clínicamente distintos. Por otra parte, si definiésemos como diferencia clínica de interés el valor 0.2, es decir, para asegurar que el Flurbiprofen tiene, o bien un 20% de éxitos, o bien, un 60%, el tamaño de muestra sería ahora

$$n = \left(\frac{1.96 \sqrt{2(0.50)(1 - 0.50)} + 1.28 \sqrt{0.40(1 - 0.40) + 0.60(1 - 0.60)}}{0.40 - 0.60} \right)^2 = 129.13$$

para $\alpha=0.05$ y $\beta=0.9$; con $\alpha=0.05$ y $1-\beta=0.8$, $z_{0.4}=0.842$ y el tamaño sería de 97 pacientes por grupo.

En resumen, cuanto más pequeños sean los dos tipos de error, es decir, más seguros estemos de los resultados, y cuanto menor sea la diferencia clínicamente importante, mayor será el tamaño de muestra necesario. El clínico que quiera llevar a cabo el estudio debe pensar muy detenidamente en estas cuestiones; en general, como valor α se suele tomar 0.05, el valor más común de β suele ser 0.20 y la diferencia clínica relevante depende de la naturaleza de cada estudio y es el clínico el único capacitado para establecerla. Es frecuente la desorientación de éste cuando pregunta a un estadístico por el tamaño de muestra de su estudio y el estadístico le responde con la pregunta acerca de la diferencia clínica en la que él está interesado; sin su respuesta no se puede conocer cuantos pacientes hay que tomar.

Si ahora nos ponemos en el lugar del lector de un artículo en que aparecen diferencias no significativas, la pregunta que nos debemos plantear: ¿cual es la potencia del test empleado?; dicho de otra manera: con el tamaño de muestra de este estudio, ¿cual es la probabilidad de obtener resultados estadísticamente significativos en caso de que los dos tratamientos sean distintos? La siguiente expresión permite el cálculo aproximado de la potencia de un test para comparar dos proporciones

$$z_{2\beta} = \sqrt{\frac{n(p_1 - p_2)^2}{2p(1 - p)}} - z_{\alpha}$$

Consideremos el estudio anterior y calculemos la potencia para detectar una diferencia de 0.1 en el porcentaje de éxitos; según la expresión anterior

$$z_{2\beta} = \sqrt{\frac{26(0.1)^2}{2(0.45)(0.55)}} - 1.96 = -1.235$$

que da lugar a una potencia de sólo 0.11, es decir un error $\beta = 0.88$; si sólo se quisiera detectar una diferencia de 0.2 todavía la potencia del test sería bastante baja, del orden del 30%. Por tanto, el hecho de que no se encontrasen diferencias en el estudio sobre el efecto del Flurbiprofen frente a la Dipirona+Hioscina es muy poco fiable, existe una gran probabilidad de aceptar que son iguales aunque se diferencien en un 0.2 de efectividad. Una amplia discusión acerca del tamaño de muestra en estudios de casos y controles puede encontrarse en Schlesselman¹⁹ y para ensayos clínicos en la monografía de Machin y Campbell²⁰.

Debido a lo anteriormente expuesto, en la fase de diseño de todo estudio es necesario el concurso de un profesional de la estadística que va a permitir conocer el tamaño mínimo de sujetos que deben participar en el estudio, evitando así los problemas que se acaban de comentar. Una de las cuestiones para evaluar un artículo a publicar en el *British Medical Journal* es precisamente si se calculó, previamente a la realización del estudio, el tamaño de muestra necesario, o bien, se han tomado los pacientes que se ha querido²¹. Nunca es una buena excusa decir que no se tienen más casos de la enfermedad que se está estudiando; si acaso fuera así, no hay más remedio que reclutar pacientes, bien esperando el tiempo suficiente, bien tomando pacientes de otros colegas. Especialmente en el caso de los ensayos clínicos controlados, no sólo es poco científico el no tomar un adecuado tamaño de muestra sino que también es un problema ético²² hacer participar a sujetos humanos en una experiencia que siempre conlleva costes y riesgos y de la que hay pocas posibilidades de obtener información fiable. Los comités hospitalarios de ensayos clínicos deberían ser muy conscientes de esta problemática.

D. TESTS PARA COMPARAR DOS GRUPOS

E. Sánchez-Cantalejo Ramírez

1. Introducción

El estudio expuesto en II.C.5 trataba de evaluar la hipótesis que establece que la probabilidad de curar con un tratamiento es igual a la probabilidad del otro tratamiento; las dos variables implicadas en el estudio eran variables dicotómicas: tipo de tratamiento y remisión del dolor cólico. Sin embargo hay algo que diferencia a estas dos variables: una es a la que podríamos llamar **variable respuesta**, la remisión del dolor, y al tratamiento le podemos denominar como **variable predictora**; lo que se ha pretendido es predecir, en base al tratamiento o variable predictora, la remisión del dolor, la variable resultado; en una terminología más estadística a la variable resultado se le denomina **variable dependiente** y a la predictora, **variable independiente**. La distinción entre los dos tipos de variables no es gratuita pues dependiendo de la naturaleza de ellas así será el análisis estadístico a realizar. Cuando la variable predictora sea dicotómica y la respuesta continua, el test a realizar es la *t* de Student; cuando tanto la predictora como la respuesta sean continuas el método es la llamada regresión lineal; el caso de una predictora continua y una respuesta dicotómica se analiza mediante la regresión logística. La Tabla II.14. resume lo que se acaba de decir.

Tabla II.14. Tipos de análisis estadístico según el tipo de variables implicadas en el estudio.

		Variable Resultado	
		Dicotómica	Continua
Predictora	Dicotómica	Test χ^2	<i>t</i> de Student
	Continua	R. logística	R. Lineal

Hay otros estudios en los que la variable resultado es el tiempo que transcurre desde la ocurrencia de un hecho hasta la ocurrencia de otro; por ejemplo, tiempo que transcurre desde el diagnóstico de cáncer de próstata hasta la muerte del paciente debida al tumor, tiempo desde la fecha del trasplante de riñón hasta la fecha de rechazo, etc. En este tipo de estudios no siempre disponemos del valor exacto de la variable resultado, el tiempo que transcurre, pues es posible que el paciente con el tumor muera de un accidente de tráfico, o que un transplantado no rechace nunca el riñón. En ese caso, estos estudios necesitan un tipo especial de análisis conocido como análisis de supervivencia. Todos estos tipos de análisis se exponen a continuación.

2. Test de la t de student

Otra forma de evaluar la diferencia entre los dos tratamientos, hubiese sido diseñar un estudio en el que la variable resultado hubiese sido el tiempo transcurrido desde la administración del tratamiento hasta la desaparición del dolor del cólico nefrítico; imaginemos que los resultados fuesen los que aparecen en la Tabla II.15. Lo que cambia respecto al estudio anterior es el carácter de la variable resultado que en este caso es el tiempo que pasa desde que se recibe el tratamiento hasta la desaparición del dolor, es decir, ahora la respuesta es una variable continua; como se acaba de decir, el método de análisis para esta situación es el test del la t de Student.

Tabla II.15. Tiempo transcurrido, en minutos, hasta la remisión del dolor cólico.

Flurbiprofen	24, 19, 30, 24, 29, 21, 26, 31, 18, 22, 31, 26, 16, 33, 21, 25, 30, 20, 19, 22, 30, 9, 17, 28, 35, 26, 23, 17, 20, 27, 18, 37
Dipirona+ Hioscina	32, 54, 36, 28, 19, 43, 28, 34, 32, 34, 29, 19, 22, 23, 33, 28, 33, 32, 49, 27, 37, 20, 28, 36, 31, 24, 20, 30, 29, 17, 28

Ahora la hipótesis nula sobre la que se quiere decidir es que el tiempo medio transcurrido hasta la desaparición del dolor es el mismo para los dos tratamientos. Los tiempos medios en las dos muestras de 32 y 31 pacientes con cólico nefrítico fué de 24.19 y 30.16 minutos para el Flurbiprofen y Dipirona+Hioscina, respectivamente,

con unas desviaciones típicas de 6.240 y 8.327 minutos. La cuestión que ahora se plantea es: ¿ la diferencia $24.19-30.16=5.97$ minutos se puede achacar al azar ?, o lo que es igual, ¿ están de acuerdo los resultados del estudio con la hipótesis de que ambos tratamientos tardan, por término medio, el mismo tiempo en hacer desaparecer el dolor? Para esta nueva situación el estadístico de contraste es

$$t_{\text{exp}} = \frac{|\bar{x}_1 - \bar{x}_2|}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

donde \bar{x}_1 y n_1 son la media y tamaño de la primera muestra, \bar{x}_2 y n_2 , los de la segunda y

$$s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$$

es una desviación típica obtenida ponderando las de las dos muestras; para nuestro ejemplo,

$$s = \sqrt{\frac{(32-1)6.24^2 + (31-1)8.327^2}{32 + 31 - 2}} = 7.34$$

por lo que

$$t_{\text{exp}} = \frac{|24.19 - 30.16|}{7.34 \sqrt{\frac{1}{32} + \frac{1}{31}}} = \frac{5.97}{1.85} = 3.22$$

Este estadístico juega un papel similar al estadístico de Pearson en el test de la χ^2 ; cuanto mayor sea su valor, más evidencias habrá de que los dos tratamientos son distintos. Para evaluar si el valor 3.22 es suficientemente grande hay que compararlo con el valor de una nueva distribución, la t de Student, distribución que depende de un parámetro que también se denomina grados de libertad cuyo valor es $n_1 + n_2 - 2$. El error mínimo al cual se puede rechazar la hipótesis de igualdad de los tiempos medios para los dos tratamientos es $P=0.002$; como este es un error muy pequeño, rechazamos tal hipótesis y concluimos que, salvo el error del 0.2%, hay evidencias para rechazar la hipótesis que establecía que los dos tratamientos tardan, por término medio, el mismo tiempo en hacer desaparecer el dolor.

Aquí son pertinentes los mismos comentarios que los realizados en II.C.5. acerca de la comparabilidad de los grupos a efecto de poder achacar solo a los distintos tratamientos los distintos tiempos medios de remisión del dolor.

Para evaluar los posibles valores de esa diferencia de tiempos medios calculamos un intervalo de confianza para esa diferencia que viene dado por la expresión

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

que para nuestro ejemplo es

$$(30.16 - 24.19) \pm 1.96(7.34) \sqrt{\frac{1}{31} + \frac{1}{32}} = (2.27, 9.67)$$

lo que nos permite decir que, con una confianza del 95%, la combinación Dipirone+Hioscina tarda, por término medio, como mínimo 2.27 y como máximo 9.67 minutos más que el Flurbiprofen en hacer desaparecer el dolor del cólico nefrítico.

Como la mayoría de los tests estadísticos, el test de Student que se acaba de realizar es aplicable solamente si se cumplen una serie de condiciones previas; para el test de Student tales condiciones establecen que las dos variables medidas, el tiempo de duración del dolor con cada tratamiento, sigan una distribución normal y, por otra parte, que las varianzas en las dos poblaciones sean parecidas. La primera se puede obviar si los tamaños muestrales sean ambos mayores que 30, o bien, que siendo menores que 30 las distribuciones de las dos variables no se diferencien mucho de la distribución normal; la segunda condición se puede resolver mediante el test de Welch²³. Otra suposición fundamental para poder aplicar este test es que las dos variables medidas sean independientes, es decir, los valores de una no vengán condicionados por los de la otra; para el caso en que no se cumpla esta condición, hay una versión del test de Student para muestras dependientes o apareadas²⁴.

Como se comentó anteriormente, a la hora de diseñar un estudio es una cuestión fundamental el plantearse el problema del tamaño de muestra. Si se establece como diferencia mínima clínicamente importante el valor δ , el tamaño de muestra a elegir en cada grupo viene dado por la expresión

$$n = \frac{(z_{\alpha} + z_{2\beta})^2 (2\sigma^2)}{\delta^2}$$

donde σ es la desviación típica común a las dos poblaciones que normalmente no se conoce y se debe estimar en base a estudios previos. Como el cálculo del tamaño de muestra para esta situación es más complejo, se debería consultar a un estadístico.

Cuando los tamaños de muestra son pequeños, en general, menores que 30 y las variables medidas no siguen una distribución normal, lo cual es bastante frecuente, no se cumplen las condiciones de aplicación del test de la t de Student. En esta situación es aplicable el llamado **test de Mann-Whitney** para muestras independientes, o bien el **test de Wilcoxon** para muestras apareadas; estos dos tests forman parte de los llamados **tests no paramétricos** o **libres de distribución**, que se pueden aplicar sin que se cumplan las condiciones exigidas al test de Student pero que tienen una menor potencia ²⁵.

Por último, poner de relieve un error muy común en las publicaciones médicas; es frecuente que en vez de evaluar dos tratamientos, se diseñe un estudio para comparar tres o más tratamientos. Para analizar esta nueva situación lo que suele hacer, indebidamente, es aplicar el test de la t de Student a cada pareja de tratamientos; este no es el camino correcto. En efecto, imaginemos un ensayo clínico en el que se quiere comparar la efectividad de cuatro formas distintas de administración de un nuevo tratamiento, frente a un tratamiento estandar; en definitiva, de lo que se trata es de comparar cinco tratamientos; si se utiliza como estrategia de análisis la comparación de todas las posibles parejas de tratamientos mediante la t de Student, tendremos que realizar un test para comparar el tratamiento 1 con el tratamiento 2, otro test para comparar el tratamiento 1 con el tratamiento 3, etc. Ya que con cinco tratamientos podemos formar diez parejas distintas de tratamientos, serán 10 el número de tests de Student a realizar. ¿Qué problemas presenta esta estrategia de análisis?; si cada uno de estos tests se hace con una confianza del 95%, es decir, un nivel de error de 0.05, se puede demostrar que siendo los cinco tratamientos igualmente efectivos, existe en el mejor de los casos una probabilidad, aproximadamente, del 30% de encontrar al menos un test significativo; en definitiva, existe una probabilidad de un 30% de rechazar la hipótesis de igualdad de los cinco tratamientos, cuando realmente lo son. Por tanto, cuando se llevan a cabo muchas comparaciones mediante tests de hipótesis, se incrementa mucho el error de rechazar la hipótesis nula siendo verdadera.

Para evitar esta dificultad conocida como el **problema de las comparaciones múltiples**, hay una solución que consiste en ser más exigente a la hora de realizar cada uno de los tests, es decir, disminuir el error α en cada uno de ellos. **La regla de Bonferroni** consiste en realizar cada test a un error α/k , donde k es el número de tests a efectuar; de esta manera el error global acumulado de todos los tests es aproximadamente el error α deseado. Para el ejemplo anterior, declararíamos cada test significativo sólo si el error de rechazo fuese como máximo $0.05/10=0.005$ para que el error global fuese 0.05 aproximadamente.

Otra alternativa para evitar el problema de las comparaciones múltiples es utilizar el llamado **Análisis de la Varianza, ANOVA (ANalysis Of VAriance)** que es un test que permite comparar globalmente tres o más tratamientos a un error prefijado, que no es más que una generalización del test t de Student. El ANOVA es un test para decidir si rechazar o no la hipótesis de igualdad de los tratamientos; si el test es significativo al error prefijado podemos decir, salvo ese error, que no todos los tratamientos son iguales, es decir, hay al menos un tratamiento que se diferencia de los restantes, pero no especifica qué tratamiento o tratamientos son los diferentes. Para contestar a esta pregunta existen una serie de tests específicos para comparaciones múltiples como son los de **Newman-Keuls, Tukey, Duncan**, etc ²⁶.

3. Regresión lineal simple

Respuestas a preguntas como las siguientes son de interés para el clínico: ¿cómo cambia el nivel de PSA cuando cambia el tamaño de la glándula prostática; ¿cómo cambia la tasa de filtración glomerular con el cambio del nivel de creatinina en sangre?, etc. En la Tabla II.16. aparecen los datos de creatinina (X) y hemoglobina (Y) medidos en 50 individuos pertenecientes a un estudio en el que se trataba de evaluar la relación entre esas dos variables en los enfermos con fallo renal crónico.

Bajo el nombre de **modelos de regresión** se engloban una serie de técnicas que tratan de medir la relación entre una variable resultado y una o unas variables predictoras; es el carácter de la variable resultado lo que determina el modelo de regresión. Como ya se avanzó, cuando la variable resultado es continua, el modelo de elección es el llamado **modelo de regresión lineal** que, dependiendo de si hay una o más predictoras se denomina **simple** o **múltiple**, respectivamente. El modelo de regresión lineal simple establece que

$$E(Y) = \beta_0 + \beta_1 X$$

donde Y es la variable resultado, X es la variable predictora y β_0 , β_1 son los parámetros del modelo; éste establece que $E(Y)$, la media de la variable resultado, es igual a una constante β_0 más el producto de otra constante β_1 por el valor de la variable predictora X ; obsérvese que en el primer miembro de la igualdad no aparece el valor de la variable resultado sino su media. Como siempre los parámetros son desconocidos y lo que haremos será, a partir de una muestra de valores de la dos variables, estimar ambos parámetros, con lo que obtendremos un modelo estimado que representaremos así

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

donde \hat{y} es la media estimada de la variable resultado y $\hat{\beta}_0, \hat{\beta}_1$ son las estimaciones de los parámetros del modelo.

Tabla II. 16. Niveles de creatinina y hemoglobina en 50 individuos con fallo renal crónico.

X Creatinina	Y Hemoglobina	\hat{y}	$Y-\hat{y}$	X Creatinina	Y Hemoglobina	\hat{y}	$Y-\hat{y}$
6.4	8.9	8.593	0.307	5.4	9.2	9.164	0.036
6.0	9.9	8.821	1.079	2.2	11.3	10.990	0.310
7.1	8.4	8.193	0.207	5.2	8.8	9.278	-0.478
4.7	9.2	9.563	-0.363	6.2	9.1	8.707	0.393
5.2	9.2	9.278	-0.078	5.0	9.0	9.392	-0.392
6.4	7.5	8.593	-1.093	5.8	8.8	8.935	-0.135
5.2	9.6	9.278	0.322	3.9	10.0	10.020	-0.020
3.6	9.3	10.191	-0.891	5.2	9.8	9.278	0.522
7.8	7.5	7.794	-0.294	5.3	9.1	9.221	-0.121
4.8	10.9	9.506	1.394	5.7	8.8	8.992	-0.192
2.5	11.2	10.819	0.381	6.8	8.6	8.365	0.235
5.8	9.4	8.935	0.465	7.3	8.2	8.079	0.121
4.6	9.2	9.620	-0.420	7.1	7.9	8.193	-0.293
7.5	8.2	7.965	0.235	3.6	10.5	10.191	0.309
5.2	9.8	9.278	0.522	6.0	9.5	8.821	0.679
5.0	8.8	9.392	-0.592	2.5	10.2	10.819	-0.619
4.3	10.4	9.792	0.608	3.5	9.2	10.248	-1.048
5.0	9.1	9.392	-0.292	6.4	8.2	8.593	-0.393
7.4	7.8	8.022	-0.222	4.8	8.7	9.506	-0.806
6.3	8.7	8.65	0.050	5.6	9.6	9.050	0.550
6.9	8.3	8.307	-0.007	4.4	9.4	9.735	-0.335
3.5	10.8	10.248	0.552	4.6	9.8	9.620	0.180
6.6	8.4	8.479	-0.079	8.5	6.7	7.394	-0.694
5.0	8.9	9.392	-0.492	3.8	9.7	10.007	-0.377
8.7	7.3	7.280	0.020	4.9	10.7	9.449	1.251

Un paso fundamental en todo análisis de regresión lineal es la representación gráfica de los datos; como quiera que el modelo establece que la relación entre la hemo-

globina y la creatinina se puede representar mediante una recta, una representación gráfica mediante una **nube de puntos** nos va a dar pistas de la adecuación de los datos observados al modelo lineal. En la Fig. II.23 aparece tal representación; según ella, no hay evidencias de que exista una relación curvilínea entre ambas variables, por lo que, en principio, el modelo de regresión lineal puede ser aplicable.

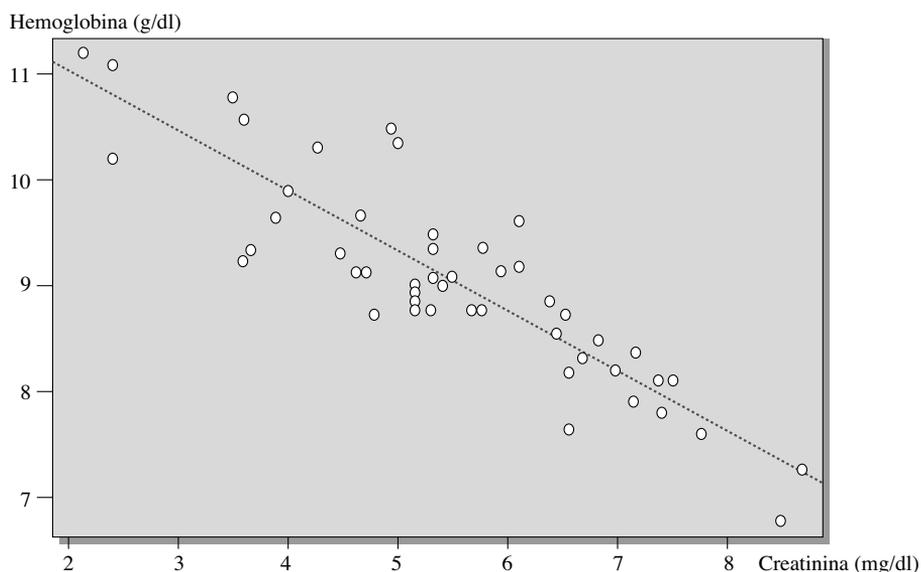


Fig. II.23. Representación gráfica de los niveles de creatinina y hemoglobina en 50 pacientes con fallo renal crónico.

Como ya se ha discutido en los tests anteriores, la única forma de conocer los parámetros del modelo sería medir ambas variables en todos los individuos de la población, en este caso todos los enfermos con fallo renal crónico. Como nuestras observaciones están basadas sólo en una muestra de 50 de tales enfermos, nos vamos a tener que conformar con estimar a tales parámetros, estimaciones que notaremos por $\hat{\beta}_0$, $\hat{\beta}_1$; el problema está ahora en elegir la recta que mejor represente el cambio en el nivel medio de la hemoglobina de estos enfermos con el cambio en el nivel de creatinina. El criterio que se emplea en regresión lineal para elegir la recta que represente a la nube de puntos de nuestra experiencia es el llamado criterio de los mínimos cuadrados; según este criterio, se elige la recta tal que la suma de los cuadrados de la diferencias entre lo observado y lo predicho por la recta sea lo más pequeña posible; de alguna manera, la recta que más se aproxime a todos los puntos de la nube. La ecuación de esta recta, estimación del modelo $E(Y)=\beta_0+\beta_1X$, la notaremos por

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

donde

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

y

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

donde x_i, y_i representan la creatinina y hemoglobina de individuo número i ; por ejemplo, $x_1=6.4$ y $y_1=8.9$.

De todas formas el cálculo es lo de menos pues hoy los ordenadores son accesibles y ellos hacen este tipo de trabajo. La tabla de resultados que arrojaría todo paquete estadístico informático con estos datos sería parecida a la Tabla II.17., en donde en una columna aparece las estimaciones de los parámetros del modelo y en otra sus respectivos errores estándar

Tabla II.17.- Estimaciones de los parámetros y sus errores estándar para los datos de la Tabla II.16.

Variable	Estimación	Error estandar
Constante	12.25	0.2984
Creatinina	-0.5708	0.05312

por lo que la recta de regresión estimada es

$$\hat{y} = 12.25 - 0.5708 \text{ creatinina}$$

cuya representación gráfica es la recta punteada de la Fig. II.23.

Veamos la interpretación de estos resultados; en primer lugar, el coeficiente β_1 de la variable predictora es conocido como **coeficiente de regresión**, y se puede interpretar como el cambio que experimenta la media de la variable resultado por unidad de aumento de la variable predictora; el cambio será aumento si el signo del coeficiente de regresión es positivo y descenso si el signo es negativo. Así, el valor -0.5708 es una estimación del cambio promedio de la hemoglobina por unidad de aumento de la creatinina; ya que este coeficiente es negativo podemos afirmar que por cada unidad, en este caso mg/dl, de aumento en el nivel de creatinina, la hemoglobina descende, por término medio, en 0.5708 g/dl y este descenso es el mismo sea cual sea el nivel de creatinina. Así, la diferencia promedio en hemoglobina entre los pacientes con 4 mg/dl de creatinina y los que tienen 5mg/dl es de 0.5708 g/dl y es la misma que entre estos últimos y los que tienen niveles de 6 mg/dl de creatinina. Una cuestión importante acerca del coeficiente de regresión es que su valor depende de las unidades de medida utilizadas para las dos variables en juego de tal manera que si no se explicitan tales unidades el coeficiente de regresión no es interpretable.

El coeficiente β_0 se denomina **término independiente** o **término constante** y es el valor de la media de la variable resultado cuando la predictora vale cero; por tanto, para nuestro ejemplo, el valor 12.25 se podría interpretar como el nivel medio de hemoglobina para los pacientes con nivel cero de creatinina, que como no tiene sentido biológico, no es interpretable. Además, este coeficiente no es interpretable en nuestro estudio porque los pacientes elegidos tienen niveles de creatinina en el rango aproximado de 2 mg/dl a 9 mg/dl y es en ese rango donde la relación con el nivel de hemoglobina es lineal. Como no sabemos si tal relación sigue la misma dinámica fuera de ese rango, el término constante no es interpretable en este estudio.

Si el coeficiente de regresión fuese cero implicaría que el modelo tomaría la forma $E(Y) = \beta_0 + (0)X = \beta_0$ es decir, la media de la variable resultado no depende de la predictora, es siempre constante, lo que equivale a decir que las dos variables son **linealmente independientes**, lo que, abusando del lenguaje, se suele traducir diciendo que son **independientes**. Debe quedar claro que con la regresión lineal lo que se estudia es la dependencia o independencia lineal y podría ocurrir que dos variables fuesen linealmente independientes pero dependientes según una relación curvilínea, lo que se puede descartar con una nube de puntos; de ahí el interés de efectuar tal representación gráfica previamente a cualquier tipo de análisis.

Aunque el coeficiente β_1 fuese cero su estimación no necesariamente tiene que serlo aunque si es esperable que sea un número próximo a cero; ¿cómo decidir si la estimación -0.5708 de nuestro ejemplo es o no cercana a cero?. Dicho de otra forma, ¿el número -0.5708 es estadísticamente distinto de cero y por tanto existe dependencia lineal entre los niveles de hemoglobina y los de creatinina en los pacientes con

fallo renal crónico? Para responder a esta pregunta no hay más que dividir el valor absoluto de la estimación del coeficiente de regresión entre su error estándar y comparar ese cociente con el valor de la distribución normal al nivel de confianza $1-\alpha$ elegido; siempre que tal cociente sea mayor que el valor de la tabla de la distribución normal se puede afirmar, salvo el error α , que las dos variables son linealmente dependientes. Según la Tabla II.17. el error estándar de la estimación del coeficiente de regresión es 0.05312; por tanto, el cociente

$$\frac{0.5708}{0.05312} = 10.74$$

hay que compararlo con el valor de la distribución normal; si se elige una confianza del 95%, el valor correspondiente de la normal es 1.96; ya que 10.74 es mayor que 1.96 podemos decir que, con un error inferior al 5%, la el nivel de hemoglobina está asociado o depende del nivel de creatinina. Si el error estándar hubiese sido 0.5312, el cociente anterior hubiese valido 1.074 que al ser menor que 1.96 no tendríamos argumentos para rechazar la independencia entre las dos variables y concluiríamos que la hemoglobina y la creatinina son linealmente independientes.

Como en los tests anteriores siempre es de interés el cálculo de un intervalo de confianza para el parámetro, en este caso el coeficiente de regresión; un intervalo de confianza para el cambio en el nivel de hemoglobina por unidad de aumento en el nivel de creatinina vendrá dado por

$$-0.5708 \pm (1.96)0.05312 = (-0.67, -0.47)$$

por lo que podemos decir con una confianza del 95% que el descenso medio de hemoglobina será como mínimo 0.47 g/dl y como máximo 0.67 g/dl, por aumento en 1 mg/dl en el nivel de creatinina. El hecho de que este intervalo de confianza no contenga el valor 0 está de acuerdo con el resultado del contraste que se acaba de realizar.

Con la ecuación de la recta antes calculada, se pueden hacer predicciones de los valores de la hemoglobina, a partir de los valores de creatinina; así, para estimar el valor medio de hemoglobina en los pacientes con un nivel de 5.2 mg/dl, no tenemos más que sustituir este valor en la ecuación antes calculada. Por tanto,

$$12.25 - 0.5708(5.2) = 9.282$$

es decir, según el modelo ajustado, los pacientes con un nivel de 5.2 mg/dl de creatinina tienen, por término medio, un nivel de 9.282 g/dl de hemoglobina; la columna encabezada por el símbolo \hat{y} de la Tabla II.16. da las estimaciones de los valores medios de la hemoglobina para los distintos valores de creatinina; según esa tabla, la estimación correspondiente a los pacientes con niveles de creatinina igual a 5.2 mg/dl es 9.278 que no coincide con el valor 9.282 arriba calculado por problemas de redondeo.

Según aparece en la Tabla II.16., se han observado cuatro pacientes con 5.2 mg/dl creatinina en los que la hemoglobina medida fue 9.6, 9.8, 8.8 y 9.8 g/dl, respectivamente; las diferencias entre cada uno de estos valores y la hemoglobina estimada para ellos

$$9.6 - 9.282, 9.8 - 9.282, 8.8 - 9.282 \text{ y } 9.8 - 9.282$$

nos dará idea de como se ajusta el modelo estimado en esos cuatro pacientes. Estas diferencias entre los valores observados y los valores predichos por el modelo \hat{y} , es decir, las cantidades $Y - \hat{y}$, son los llamados **residuales** cuyos valores también aparecen en la Tabla II.16. El estudio de los residuales es una parte muy importante, aunque muy olvidada, del análisis de regresión; de igual manera a lo comentado en los dos tests anteriores, el test de la χ^2 y el test de la t de Student, el análisis de regresión lineal se puede llevar a cabo solo si se cumplen ciertas condiciones que de no darse invalidan las conclusiones que se puedan derivar de él. Tales condiciones, que escapan al nivel de esta monografía, se comprueban precisamente a partir del análisis de los residuales así como otras cuestiones de interés como son la identificación de las llamadas **observaciones extremas (outliers)** que son aquellas que tienen residuales grandes, es decir, hay gran desacuerdo entre lo observado y lo predicho por el modelo, y las denominadas **observaciones influyentes** que son aquellas cuya presencia o ausencia cambia sensiblemente las estimaciones de los parámetros del modelo. Debido a la complejidad de este tipo de análisis de los residuales, el clínico debería buscar la colaboración de un estadístico.

4. Correlación

Además de estudiar la existencia de una asociación lineal entre el nivel de hemoglobina y creatinina en enfermos con fallo renal crónico, en muchas ocasiones es de interés responder a preguntas parecida a la siguiente: ¿cómo de fuerte es esa asociación entre la hemoglobina y la creatinina? Dicho de otro modo, conocido el nivel de creatinina de un enfermo, ¿cómo de precisa es la estimación, basada en el modelo anterior, de su nivel de hemoglobina? Con el análisis de regresión lineal se evalúa el

cambio que experimenta la media de la variable resultado con el aumento en una unidad de la predictora, cambio que se cuantifica mediante el coeficiente de regresión. Sin embargo, en la Fig. II.24 aparecen dos nubes de puntos que dan lugar al mismo coeficiente de regresión pero es evidente que la relación que representan las dos gráficas es distinta; en la de la izquierda la relación entre las dos variables es más estrecha que en la de la derecha; por tanto, el coeficiente de regresión no capta la mayor o menor **fuerza de la asociación** entre dos variables. El **coeficiente de correlación de Pearson** entre dos variables continuas es una medida de la fuerza de la asociación entre tales variables; de alguna manera, similar al riesgo relativo para el caso de variables categóricas. El coeficiente de correlación entre dos variables puede ser cualquier número comprendido entre -1 y 1; el signo de este coeficiente indica el sentido del cambio de la variable resultado con el cambio de la predictora; así, un coeficiente de correlación negativo indica que a mayor valor de la predictora, menor valor de la variable resultado; si el signo es positivo indica que cuando aumenta una, también aumenta la otra. El valor absoluto del coeficiente de correlación informa de la fuerza de la asociación entre las dos variables; un coeficiente de correlación con valor absoluto próximo a 1, es decir, un coeficiente de correlación próximo a -1 o a 1, indica que las dos variables están estrechamente asociadas. Un coeficiente de correlación próximo a cero indica poca o nula asociación entre las variables.

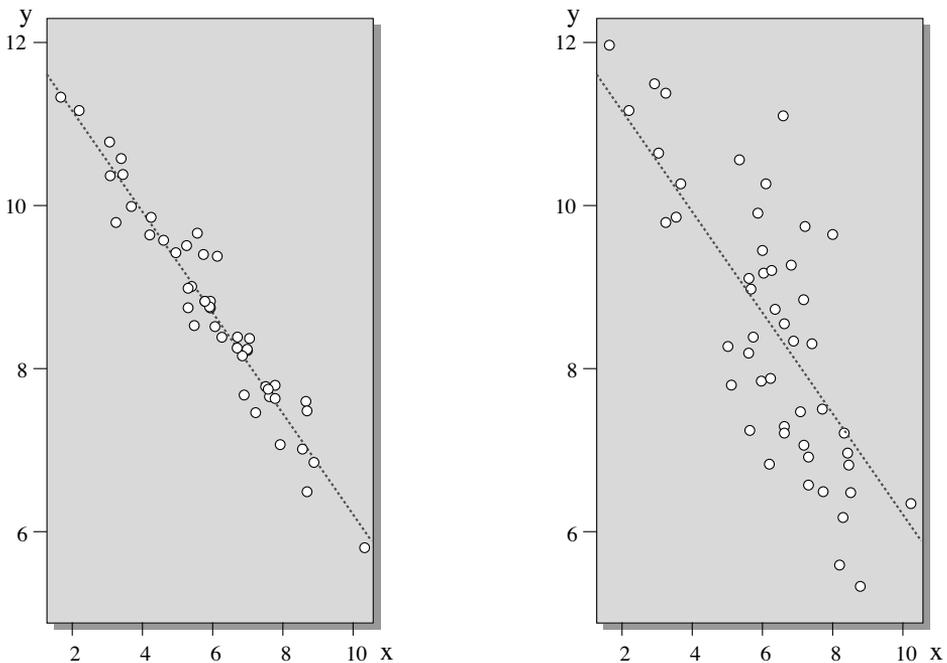


Fig. II.24.- Nubes de puntos con distinta fuerza de asociación.

Como todo parámetro no va a ser conocido, por lo que hay que conformarse con una estimación a partir de los datos de la muestra elegida; tal estimación viene dada por la expresión

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Una estimación del coeficiente de correlación calculada a partir de la nube de puntos de la gráfica de la izquierda de la Fig. II.24 es -0.98, mientras que para la de la derecha es -0.72; el signo negativo de ambos viene reflejado por el descenso de la variable resultado con el aumento de la predictora; el valor absoluto de la primera estimación, 0.98, es mayor que el de la segunda, 0.72, lo que significa que las variables representadas en la gráfica de la izquierda tienen una asociación entre ellas más fuerte que las de la gráfica de la derecha.

Para el ejemplo que de hemoglobina y la creatinina,

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = -60.89, \quad \sum (x_i - \bar{x})^2 = 106.67 \quad \text{y} \quad \sum (y_i - \bar{y})^2 = 49.205$$

por tanto, la estimación del coeficiente de correlación es

$$r = \frac{-60.89}{\sqrt{(106.67)(49.205)}} = -0.84$$

Ya que el valor 0 del coeficiente de correlación indica no asociación entre las dos variables, la creatinina y la hemoglobina, la cuestión que ahora se plantea es de decidir si la estimación -0.84 es estadísticamente distinta de cero, pues cero es el valor del coeficiente de correlación que establece la no asociación o independencia lineal entre las dos variables. Para ello no hay más que calcular el valor del estadístico

$$\frac{|r|}{\sqrt{\frac{(1 - r^2)}{n - 2}}}$$

donde $|r|$ es el valor absoluto del coeficiente de correlación estimado y n es el número de individuos en los que se han medido ambas variables; por tanto

$$\frac{0.84}{\sqrt{\frac{(1 - 0.84^2)}{50 - 2}}} = 10.73$$

que comparado con la valor de la distribución t de Student con $50-2=48$ grados de libertad, permite rechazar la hipótesis de no asociación con una $P < 0.0001$. Por tanto, con un error menor que éste podemos afirmar que existe asociación entre las dos variables y el grado de de tal asociación viene cuantificada por el valor -0.84 .

Al contrario que el coeficiente de regresión, el coeficiente de correlación no depende de las unidades de medida en que se expresen las variables aunque entre los dos coeficientes existe una relación dada por la expresión

$$r = \hat{\beta}_1 \frac{s_x}{s_y}$$

por tanto que el coeficiente de correlación sea cero, lo que equivale a afirmar que las dos variables son independientes, implica que el coeficiente de regresión también vale cero; de todas formas, ambos coeficientes expresan conceptos completamente distintos.

El cuadrado del coeficiente de correlación se conoce con el nombre de **coeficiente de determinación** que podemos estimar por la cantidad r^2 . Este coeficiente de determinación mide la parte de variabilidad de la variable resultado que puede ser explicada por la variable predictora; para nuestro ejemplo, una estimación del coeficiente de determinación es $0.84^2=0.71$, de lo que se puede afirmar que el 71% de la variabilidad de las hemoglobinas de los 50 pacientes puede ser explicada por la distinta creatinina de ellos; resta por explicar un 29% que es achacable a otras variables distintas a la creatinina. De todas formas, este coeficiente de determinación cobra su máximo interés en el caso de la regresión lineal multivariante que más adelante trataremos.

Una alternativa al coeficiente de correlación de Pearson

Como en todos los tests vistos hasta ahora, señalar que para poder hacer inferencias acerca del coeficiente de correlación es necesario que al menos una de las variables medidas siga una distribución Normal. Hay muchas ocasiones en que interesa medir la fuerza de la asociación entre dos variables que ninguna de ellas cumple la condición de seguir una Normal; por ejemplo, cuando las variables son ordinales o

cuantitativas discretas, tal condición no se cumple por lo que no es válido el análisis anterior. En este caso hay una alternativa conocida como el **coeficiente de correlación de Spearman** que se puede utilizar aunque ninguna de las variables siga una Normal. El coeficiente de correlación de Spearman se basa en asignar **rangos** a cada valor de cada una de las variables; veamos como se lleva a cabo tal asignación. Consideremos, por ejemplo, los valores medidos de la creatinina; de lo que se trata es de ordenar los valores de la creatinina en forma creciente y asignarle un 1 al valor más pequeño de esta variable que es 2.2 mg/dl; al valor de creatinina inmediatamente superior se le asignará un 2, al siguiente un 3, etc. En ocasiones se pueden producir lo que se denomina un **empate**, es decir, hay dos o más valores iguales; en ese caso se le asigna a cada uno la media aritmética de los rangos que les corresponderían; por ejemplo, el valor inmediatamente superior a 2.2 es 2.5 mg/dl, y hay dos individuos que tienen tal nivel de creatinina; a uno de ellos le correspondería el rango 2 y al otro el 3; lo que vamos a hacer es asignar a cada uno de ellos el valor $(2+3)/2=2.5$. En la Tabla II.18. aparecen los rangos de las dos variables en los 50 individuos.

Tabla II.18.- Rangos asignados a los valores de creatinina y hemoglobina.

Creatinina (Rangos)	38 33.5 43.5 14 24 38 24 6.5 48 15.5 2.5 31.5 12.5 47 24 19.5 10 19.5 46 36 42 4.5 40 19.5 50 28 1 24 35 19.5 31.5 9 24 27 30 41 45 43.5 6.5 33.5 2.5 4.5 38 15.5 29 11 12.5 49 8 17
Hemoglobina (Rangos)	20.5 41 11.5 28 28 3.5 35.5 31 3.5 48 49 32.5 28 8 39 17.5 44 24 5 14.5 10 47 11.5 20.5 2 28 50 17.5 24 22 17.5 42 39 24 17.5 13 8 6 45 34 43 28 8 14.5 35.5 32.5 39 1 37 46

Pues bien, el coeficiente de Spearman se calcula igual que el coeficiente de correlación con la diferencia de que en lugar de utilizar para cada individuo las parejas de valores de creatinina y hemoglobina, se utilizan sus rangos correspondientes; otra manera equivalente de cálculo del coeficiente de Spearman es

$$r_s = 1 - \frac{6 \sum d_i^2}{n^3 - n}$$

donde d_i es la diferencia entre los dos rangos asociados al individuo i ; es decir,

$$r_s = 1 - \frac{6 [(38-20.5)^2 + (33.5-41)^2 + \dots + (17-46)^2]}{50^3 - 50} = -0.81$$

Como ocurre en este ejemplo, es frecuente que el coeficiente de correlación de Pearson y el de Spearman tengan valores parecidos, en nuestro ejemplo -0.84 y -0.81 , respectivamente; de todas formas algunos autores abogan por el mayor uso del segundo pues, además de poder calcularse en cualquier situación, tiene como ventaja sobre el de Pearson el hecho de que detecta asociaciones distintas de la lineal. Con muestras de tamaño mayor que 30, el test para el coeficiente de Spearman se puede llevar a cabo como se ha hecho en el caso del coeficiente de Pearson.

Por último, señalar que tanto el análisis de regresión como de correlación tan solo puede poner en evidencia asociaciones entre variables; para dar el salto de la asociación a la causalidad se requieren otros argumentados adicionales que se discutirán en el capítulo II.E.

5. Regresión logística

El análisis de regresión lineal que se acaba de exponer utilizaba como variable resultado una variable continua. Sin embargo, es muy frecuente encontrarse con estudios en donde la variable resultado es la presencia o ausencia de alguna característica de interés en los sujetos objeto de estudio; ¿está ligada la presencia o no de hipertensión en los individuos al nivel de colesterol en sangre?; ¿qué relación existe entre el desarrollo o no de cáncer de esófago con la ingesta de alcohol? etc. Todas estas preguntas tienen en común el hecho de que la variable resultado es dicotómica y lo que se quiere es estudiar la relación entre ésta y una variable predictora, que puede ser de cualquier tipo: categórica o cuantitativa. El modelo de regresión para estudiar este tipo de relaciones se conoce con el nombre de **regresión logística simple**.

Para analizar este tipo de relaciones lo que trataremos será de expresar la probabilidad de que un individuo presente la característica de interés, desarrollar un cáncer, ser hipertenso, etc., en función de la variable predictora; si notamos por P tal probabilidad, siguiendo lo visto en la regresión lineal, podríamos expresar el modelo de la forma

$$P = \beta_0 + \beta_1 X$$

donde, como antes, X es la variable predictora. Sin embargo, el modelo este tiene algunos inconvenientes que lo invalidan para nuestro propósito; una de las razones para la no utilidad de este modelo es que podría ocurrir que calculadas las estimaciones de los parámetros del modelo, la sustitución de algunos valores de la predictora dieran lugar a valores no permitidos para una probabilidad, es decir, valores fuera del inter-

valo (0,1). Para evitar ese problema lo que vamos a hacer es realizar una transformación de la probabilidad de presentar la característica de interés, denominada **transformación logit**, que consiste en el logaritmo de la ventaja $P/(1-P)$ de presentar la característica, por lo que el modelo lo formularemos así

$$\log \frac{P}{1-P} = \beta_0 + \beta_1 X$$

que no presenta los problemas antes comentados. A $\log[P/(1-P)]$ se le llama **logit(P)**.

Recuérdese que en el caso de la regresión lineal, el coeficiente β_1 medía el cambio de la media de la variable resultado por unidad de aumento de la predictora. ¿Como interpretar β_1 en el caso de la regresión logística? Para ello supongamos dos individuos con valores x_1 y x_2 de la predictora; si notamos por P_1 y P_2 las probabilidades de que esos individuos presenten la característica, tendremos que

$$\log \frac{P_1}{1-P_1} = \beta_0 + \beta_1 x_1 \qquad \log \frac{P_2}{1-P_2} = \beta_0 + \beta_1 x_2$$

y restando esas dos igualdades

$$\log \frac{P_1}{1-P_1} - \log \frac{P_2}{1-P_2} = \log \frac{\frac{P_1}{1-P_1}}{\frac{P_2}{1-P_2}} = \beta_1(x_1 - x_2)$$

Si el valor x_1 es una unidad mayor que el valor x_2 , y ya que

$$\frac{\frac{P_1}{1-P_1}}{\frac{P_2}{1-P_2}} = O.R.$$

es la *odds ratio*, podemos afirmar que para dos individuos que se diferencian en una unidad en relación a la predictora

$$\log(O.R.) = \beta_1$$

o lo que es igual

$$O.R. = e^{\beta_1}$$

donde e es el número que sirve de base a los logaritmos neperianos, aproximadamente, 2.72. Por tanto, en el modelo de regresión logística β_1 es el logaritmo de la O.R. entre dos individuos que se diferencian en una unidad en términos de la variable predictora, o bien, elevando e a β_1 se obtiene el valor de la O.R. entre esos dos individuos. En el caso en que $\beta_1=0$, ello implica que el $\text{logit}(P)=\beta_0+(0)X=0$, es decir, no cambia con X , o lo que es igual, $O.R.=e^0=1$, lo que indica que las dos variables son independientes, no existe asociación.

En la Tabla II.19. se presentan unos datos hipotéticos acerca de la relación entre la edad gestacional y la mortalidad en recién nacidos; para ello imaginemos un estudio de casos y controles con los siguientes resultados adaptados de Brand²⁷.

Tabla II.19.- Mortalidad y edad gestacional, en semanas, en 410 recién nacidos.

		Edad gestacional						
		24	25	26	27	28	29	Total
Vivo		3	16	38	49	45	30	181
Muerto		26	56	69	50	21	7	229

El problema está en estudiar como cambia el riesgo de muerte con la edad gestacional; para ello tenemos que estimar los parámetros del modelo a partir de los datos del estudio. El criterio de estimación de los parámetros de un modelo de regresión logística, denominado el **método de la máxima verosimilitud** es bastante más complejo que el de los mínimos cuadrados empleado en el caso de la regresión simple, pues no hay fórmulas explícitas para las estimaciones de los parámetros como ocurría en la regresión lineal; el cálculo de las estimaciones se basa en procesos matemáticos complejos que requieren gran cantidad de cálculo hasta tal punto que solo son posibles mediante el auxilio de un ordenador.

La salida de cualquier paquete informático es parecida a esta

Tabla II.20.- Estimaciones de los parámetros del modelo logístico para los datos de la Tabla II.19.

Variable	Estimación	Error estandar
Constante	18.36	2.375
Edad gestacional	-0.6819	0.08908

Para ver si existe asociación entre la mortalidad infantil y la edad gestacional, es decir, para decir si la estimación -0.6819 es estadísticamente distinta de cero, no hay más que dividir tal estimación por su error estandar

$$0.6819 / 0.08908 = 7.65$$

y compararlo con el valor de la distribución Normal al nivel de confianza elegido; para este caso, el error α es despreciable, por lo que salvo ese error, podemos afirmar la existencia de una asociación entre la mortalidad de los recién nacidos y su edad gestacional; evidentemente en el rango de 24 a 29 semanas estudiado.

Si notamos por P la probabilidad de que muera un recién nacido, una estimación del modelo viene dada por la ecuación

$$\log \frac{P}{1 - P} = 18.36 - 0.6819 \text{ Edad}$$

donde -0.6819 es la estimación de β_1 , el **coeficiente de regresión logística**. El signo negativo del coeficiente nos indica que a mayor edad del recién nacido, menor logaritmo de la «ventaja» de morir, en definitiva, menor riesgo de muerte. Elevando el número e a este valor nos da la estimación de la O.R., el riesgo, de morir de un recién nacido respecto de otro con una unidad, semana, menos. Por tanto, $e^{-0.6819} = 0.506$ es el riesgo de morir de un recién nacido respecto de otro con una semana menos; dicho de otro modo, un recién nacido, en este rango de edad gestacional, con una edad determinada entre 24 y 29 semanas, está a $1/0.506=1.98$ veces más riesgo de morir que un recién nacido con una semana más de edad gestacional.

¿ Como es el riesgo de morir de un recién nacido con 25 semanas en relación a uno con 29 semanas?; si el primero está a 1.98 veces más riesgo que uno de 26 sema-

nas y éste a su vez a 1.98 veces más que uno de 27 semanas, etc., el riesgo del de 25 semanas será $(1.98)(1.98)(1.98)(1.98)=15.34$ veces mayor que el de 29 semanas. En general, el riesgo de un recién nacido de x_1 semanas respecto a otro de x_2 semanas viene dado por la expresión

$$e^{-0.6931(x_1 - x_2)}$$

que para nuestro ejemplo es

$$e^{-0.6819(25-29)} = e^{-0.6819(-4)} = 15.30$$

donde la diferencia entre este valor y el 15.34 antes calculado se debe a problemas de redondeo. En definitiva, un recién nacido de 25 semanas está a un riesgo de morir 15 veces mayor que un recién nacido de 29 semanas. La Fig. II.25 ilustra los logits observados y los ajustados mediante el modelo de regresión logística.

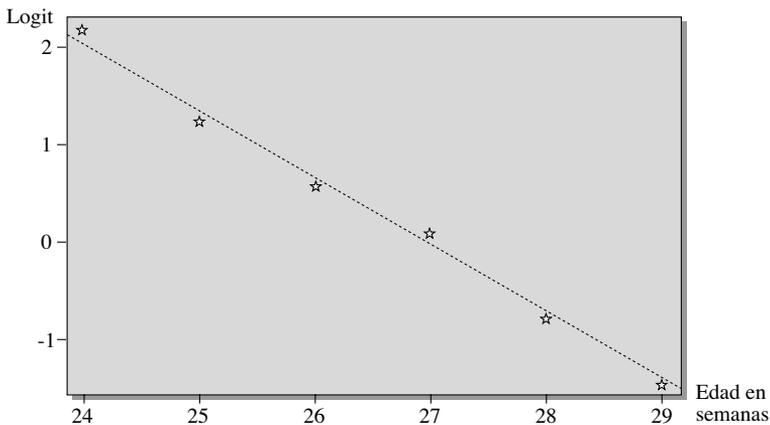


Fig. II.25.- Representación de los logits en las diferentes edades.

La interpretación del coeficiente β_0 no la comentamos ya que depende del tipo de estudio; en los diseños de casos y controles no es interpretable si no se dispone de otra información que no suele estar disponible; de todas formas, nuestro interés estaba centrado en la asociación entre mortalidad y edad gestacional y es el coeficiente β_1 el que mide tal asociación.

Como en la regresión lineal, el modelo logístico no se puede aplicar sin más pues el modelo establece que los $\text{logit}(P)$ deben seguir una recta; en caso contrario, habría

que ajustar un modelo más complicado. Por tanto, siempre es necesario representar gráficamente los *logit* (P) y la predictora para evaluar su linealidad; cuando los datos no son agrupados, existen procedimientos, que no comentaremos por su complejidad, que permiten hacerse una idea de la linealidad.

A partir de las estimaciones anteriores se puede construir un intervalo de confianza para la O.R.; así, un intervalo al 95% de confianza para el logaritmo de la O.R. viene dado por

$$-0.6819 \pm 1.96 (0.08908) = (-9.856, -0.507)$$

por lo que un intervalo de confianza para la O.R. es

$$(e^{-0.856}, e^{-0.507}) = (0.42, 0.60)$$

por tanto, la O.R. de morir de un recién nacido respecto de otro con una semana menos está entre 0.42 y 0.60; dicho de otra forma, un recién nacido está a un riesgo de morir como mínimo de $1/0.60=1.67$ veces y como máximo de $1/0.42=2.38$ veces que un recién nacido de una semana más de edad gestacional. Como en regresión lineal, las estimaciones de los parámetros dependen de las unidades de medida, en este caso semanas; si la edad gestacional se hubiese medido en días, el lector puede comprobar que el coeficiente estimado hubiera sido $-0.6819/7=-0.0974$.

6. Análisis de supervivencia

En ocasiones la investigación clínica está orientada hacia estudios de seguimiento en pacientes afectados de cáncer con el fin de conocer el tiempo de supervivencia, el porcentaje de individuos que sobreviven a una fecha determinada o el número de sujetos que mueren en un periodo de tiempo dado. Cuando se tienen dos o más grupos de pacientes con características diferentes determinadas por la edad, sexo, estadio del tumor, tratamiento aplicado o cualquier otra variable, el interés puede radicar en comparar los tiempos de vida de los distintos grupos y decidir cual de ellos tiene una longevidad mayor, o si por el contrario los tiempos de vida son parecidos en todos ellos.

Las técnicas estadísticas que resuelven los problemas anteriores constituyen lo que se conoce por **análisis de supervivencia**; éstas son técnicas cuyo objetivo es estudiar el tiempo que transcurre desde la ocurrencia de un determinado suceso: comienzo de un tratamiento, diagnóstico de cáncer, remisión del cáncer, etc., hasta la ocurrencia de otro: curación de la enfermedad, muerte, ocurrencia de recidiva, etc., respectivamente. El nombre de estas metodologías se debe al hecho de que se han empleado hasta

la fecha principalmente para el estudio del tiempo desde el diagnóstico de cáncer hasta la muerte del paciente.

Obviamente es imposible observar y estudiar la evolución de los individuos que forman la cohorte indefinidamente, por lo que la mayoría de los estudios de seguimiento tienen una fecha de comienzo y otra de finalización del estudio que marcan el periodo de tiempo durante el que los pacientes son observados, donde, generalmente, la fecha de inicio coincide con el diagnóstico del tumor y la fecha de finalización está predeterminada por el investigador dependiendo del diseño realizado o de los datos disponibles. Durante este periodo de tiempo el seguimiento es individual, pudiendo ocurrir que el paciente fallezca, se pierda el contacto con él en un instante determinado o alcance con vida la fecha considerada como final del estudio; las dos últimas situaciones representan lo que se denominan datos **censurados**, es decir, individuos que no han fallecido y por tanto no aportan información al estudio sobre el tiempo exacto de vida pero que no pueden ser eliminados porque sabemos que el tiempo de vida es superior a la fecha donde se perdió el contacto o se finalizó el seguimiento. La Tabla II.21. presenta los tiempos de supervivencia de un grupo de 10 pacientes diagnosticados de cáncer de próstata observados durante un periodo de 10 años, 1983-1992.

Tabla II.21.- Tiempos de supervivencia en 10 pacientes con cáncer de próstata.

Paciente nº	Tiempo de Supervivencia (en años)	Censura
1	10.00	0
2	6.51	1
3	7.27	0
4	4.82	1
5	8.00	0
6	9.00	0
7	7.70	1
8	3.44	1
9	1.56	1
10	0.65	1

La Fig. II.26 muestra la representación gráfica de la experiencia de supervivencia de este grupo de enfermos; el eje horizontal de esta gráfica corresponde a lo que se denomina año calendario o tiempo de seguimiento, es decir, el año cero es el inicio del estudio, 1983, y el 10 el final del seguimiento, 1992; en esta situación cada individuo entra en el estudio en un instante diferente y el cálculo del **tiempo de supervivencia** ha de hacerse restando el tiempo de salida (bien por muerte, bien por cen-

sura), correspondiente al extremo superior del segmento, menos el tiempo de entrada, correspondiente al extremo inferior.

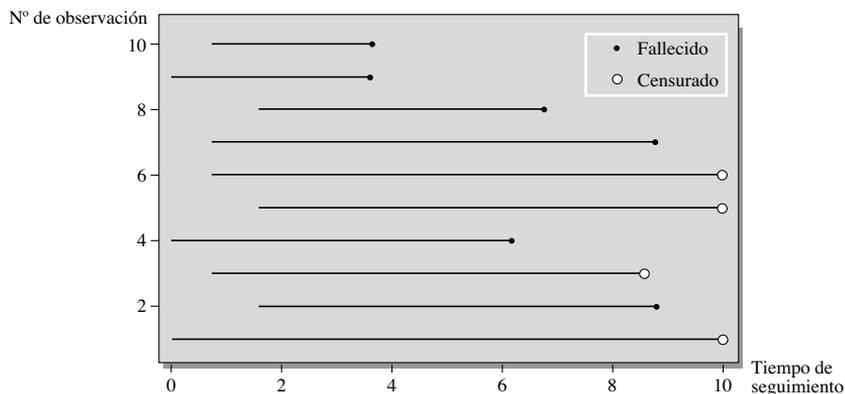


Fig. II.26.- Representación gráfica de los tiempos de entrada y salida del seguimiento en diez pacientes con cáncer de próstata.

Mucho más sencillo es considerar en el eje horizontal el tiempo de supervivencia para cada paciente, con lo que todos entrarían en el estudio al tiempo cero y el seguimiento quedaría gráficamente de la siguiente forma

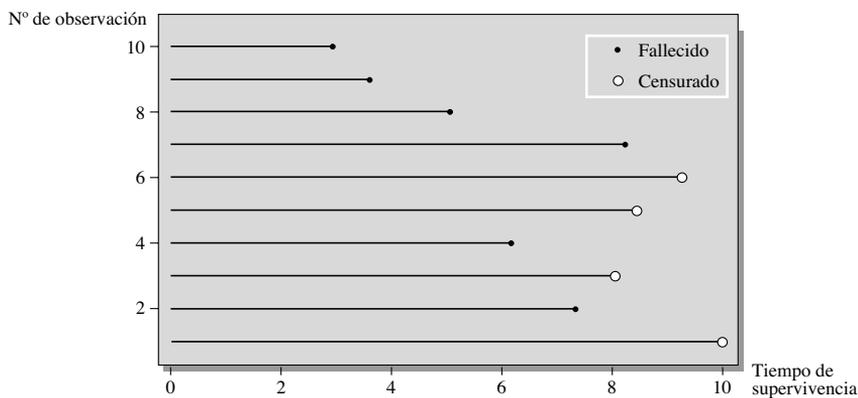


Fig. II.27.- Representación gráfica de los tiempos de supervivencia de diez pacientes con cáncer de próstata.

Así, para realizar cualquier estudio de supervivencia tendremos que disponer de, al menos, la variable tiempo de vida o supervivencia de cada paciente y si es o no una observación censurada, codificando ésta usualmente como 1 si ha fallecido durante el seguimiento o 0 si es censurado.

Pretendemos ahora calcular la probabilidad de que un paciente afectado de cáncer de próstata sobreviva a un tiempo t determinado (1 año, 5 años, etc.), a partir del

diagnóstico de su enfermedad, es decir, buscamos una curva de supervivencia donde en el eje horizontal esté representada la variable Tiempo y en el eje vertical la probabilidad de que un individuo de estas características sobreviva a un instante de tiempo dado. Usualmente esta curva suele estar representada por

$$S(t) = P[T > t] \quad t \geq 0$$

donde $P[T > t]$ indica la probabilidad de sobrevivir más de un tiempo t . Como siempre, de entrada vamos a renunciar a conocer la supervivencia poblacional y lo que trataremos será de estimarla en base a la supervivencia de nuestros pacientes, para desde ella hacer inferencias a la población de los pacientes de cáncer de próstata. Hay distintos procedimientos de estimar la supervivencia; a continuación se presenta el más utilizado en la investigación sanitaria.

6. 1. El método de Kaplan-Meier

Uno de los métodos más populares para construir estas curvas de supervivencia es el de Kaplan-Meier que da una metodología sencilla para el cálculo de las probabilidades de vida, incluso cuando hay datos censurados. Para ello ordenaremos los tiempos de supervivencia de la Tabla II.21. de menor a mayor para tener una secuencia de lo que ha ocurrido durante el seguimiento, como aparece en la Tabla II.22.

Tabla II.22.- *Tiempos de supervivencia ordenados de 10 pacientes.*

Tiempo de supervivencia (En años)	Censura
0.65	1
1.56	1
3.44	1
4.82	1
6.51	1
7.27	0
7.70	1
8.00	0
9.00	0
10.00	0

De esta forma se tiene que, de los 10 pacientes que constituyen la cohorte, únicamente han fallecido 6 en los instantes de tiempo $t_1=0.65$, $t_2=1.56$, $t_3=3.44$, $t_4=4.82$, $t_5=6.51$ y $t_6=7.70$ y el resto han sido censurados, es decir, la observación 3 tiene un tiempo de vida superior a 7.27 años y los restantes pacientes sobrevivieron al menos 8 años. Llegados a esta situación podemos definir para cada instante t_k , $k=1, \dots, 6$, lo que se denominan grupos de riesgo R_k , constituidos por el número de individuos que

permanecen con vida justo antes de dicho instante; así, inmediatamente antes del primer tiempo de muerte, los 0.65 años, los 10 pacientes están a riesgo de muerte por lo que $R_1=10$; sólo 9 de los 10 pacientes están a riesgo de muerte inmediatamente antes de los 1.56 años, pues uno de los pacientes ya murió; por tanto, $R_2=9$; de forma similar, los restantes grupos de riesgo para los diferentes tiempos de muerte son $R_3=8$, $R_4=7$, $R_5=6$ y $R_6=4$.

Ya que la primera muerte no se produce hasta el tiempo $t_1=0.65$, tendremos que de 10 individuos que comienzan el estudio sobreviven 10 en el instante de tiempo 0, y por tanto la probabilidad de sobrevivir más de 0 años es $10/10=1$, es decir, $\hat{S}(0)=P[T>0]=1$, donde $\hat{S}(t)$ representa la estimación de la curva de supervivencia en el tiempo t . Ya que desde el inicio del estudio hasta los 0.65 años no se produce ninguna muerte, en este periodo la supervivencia seguirá siendo la misma, es decir, $\hat{S}(t)=1$ en todo instante de tiempo inferior a 0.65 años, o sea, $0 < t < 0.65$.

En el tiempo $t_1 = 0.65$ muere uno de los pacientes de los 10 que estaban a riesgo de morir, el grupo R_1 , por lo que quedan 9 con vida, luego la probabilidad de sobrevivir al tiempo $t_1=0.65$ es el cociente $\hat{S}(0.65)= 9/10 = 0.9$. Entre los nueve que quedan a riesgo de muerte después del tiempo 0.65 años, el grupo R_2 , no se produce ninguna muerte hasta los 1.56 años; así, entre los 0.65 y los 1.56 años, la estimación de la supervivencia sigue siendo 0.9. En el tiempo 1.56 se produce una muerte entre los 9 a riesgo, por lo que la probabilidad de sobrevivir, de entre los que sobrevivieron al primer tiempo de muerte, al tiempo 1.56 será $8/9$.

Los pacientes que continúan con vida después del tiempo $t_2=1.56$ son aquellos que sobrevivieron a $t_1=0.65$ y a $t_2=1.56$, luego la probabilidad de sobrevivir a $t_2=1.56$ será el producto de la probabilidad de sobrevivir a $t_1=0.65$ por la de sobrevivir a $t_2=1.56$, es decir, $\hat{S}(1.56)= (9/10) (8/9)= 0.8$. De igual forma, $\hat{S}(3.44)$ es la probabilidad de superar con vida el instante de tiempo $t_1=0.65$, $t_2=1.56$ y $t_3=3.44$, luego será el producto de tres probabilidades dado por $\hat{S}(1.56)=(9/10) (8/9)(7/8)=0.7$.

En general, el cálculo de las probabilidades de vida en cada instante se consigue mediante la expresión

$$\hat{S}(t_k) = \frac{R_1 - d_1}{R_1} \times \dots \times \frac{R_k - d_k}{R_k}$$

donde k es el número de tiempos en los que se ha producido una muerte, siendo d_1 el número de muertes ocurridas en $t_1=0.65$, d_2 el número de muertes ocurridas en $t_2=1.56$, etc., que en este caso el número de muertes es igual a 1 en todos los tiempos. Así, para el tiempo $t_3=3.44$,

$$\hat{S}(3.44) = \frac{10 - 1}{10} \times \frac{9 - 1}{9} \times \frac{8 - 1}{8} = 0.7$$

Esta forma de estimar la supervivencia es el llamado método de Kaplan-Meier, que nos permite confeccionar la Tabla II.23. donde aparecen las probabilidades de supervivencia.

Tabla II.23.- Estimación de la supervivencia en 10 pacientes con cáncer de próstata.

Tiempos de muerte (en años)	Personas a riesgo R_k	Número de muertes d_k	Supervivencia $\hat{S}(t_k)$
0.65	10	1	0.900
1.56	9	1	0.800
3.44	8	1	0.700
4.82	7	1	0.600
6.51	6	1	0.500
7.70	4	1	0.375

Obsérvese que entre dos instantes de tiempo en que ocurre una muerte, la supervivencia se mantiene constante puesto que en ese periodo de tiempo no fallece ningún paciente; así, si $\hat{S}(0.65)=0.9$ y $\hat{S}(1.56)=0.8$, la supervivencia entre 0.65 y 1.56 se mantiene igual a 0.9 por ser el número de personas que continúan con vida el mismo en ese periodo de tiempo. Igualmente, si $\hat{S}(6.51)=0.5$ y $\hat{S}(7.70)=0.375$, el tiempo de supervivencia entre 6.51 y 7.70 es 0.5 puesto que entre ambos instantes de tiempo no se produce ningún fallecimiento. Teniendo en cuenta estas consideraciones puede construirse una curva de supervivencia escalonada, representando en el eje horizontal el tiempo en años y en el vertical las probabilidades de vida, obteniéndose la Fig. II.28

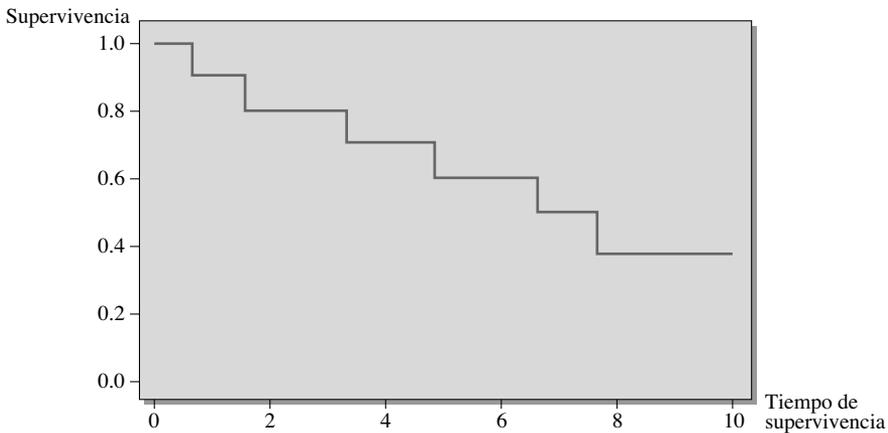


Fig. II.28.- Curva de supervivencia estimada por el método Kaplan-Meier.

En ella puede verse que la probabilidad de continuar con vida tras 4 años a partir del diagnóstico del tumor es 0.7, es decir, el 70% de la población tiene una supervivencia superior a 4 años, vive más de 4 años, tras el diagnóstico de cáncer de próstata. Como se puede ver, la representación gráfica de la curva de supervivencia es una línea en forma de escalera, donde cada peldaño viene condicionado por la presencia de alguna muerte en el tiempo correspondiente y la supervivencia se mantiene constante entre cada dos tiempos consecutivos de muerte.

6. 2. Comparación de curvas de supervivencia. El test log-rank

Una de las aplicaciones más interesantes del análisis de supervivencia está en comparar las curvas de supervivencia de dos o más grupos de pacientes para contrastar si existen diferencias significativas entre ellas, lo que puede hacerse a través de uno de los test de hipótesis más usados en investigaciones clínicas: el test de Mantel-Haenszel o Log-Rank. Para ver cómo se construye este contraste de hipótesis utilizaremos los datos de la Tabla II.24. que representan los tiempos de supervivencia en días de dos grupos de pacientes diagnosticados con cáncer y diferenciados por la variable Sexo, donde el símbolo * indica que la observación ha sido censurada.

Tabla II.24.- Tiempos de supervivencia, expresados en días, en dos grupos de pacientes con cáncer.

GRUPO 1 (Hombres)	8, 78*, 109, 109, 137, 196, 253, 325, 345, 560, 730
GRUPO 2 (Mujeres)	82, 253, 342, 381, 404, 662, 838, 1114, 1159

Las curvas de supervivencia estimadas mediante el método de Kaplan-Meier para ambos grupos están representadas en la Fig. II.29, donde, aparentemente, puede apreciarse una clara diferencia entre hombres y mujeres, con ventaja para éstas.

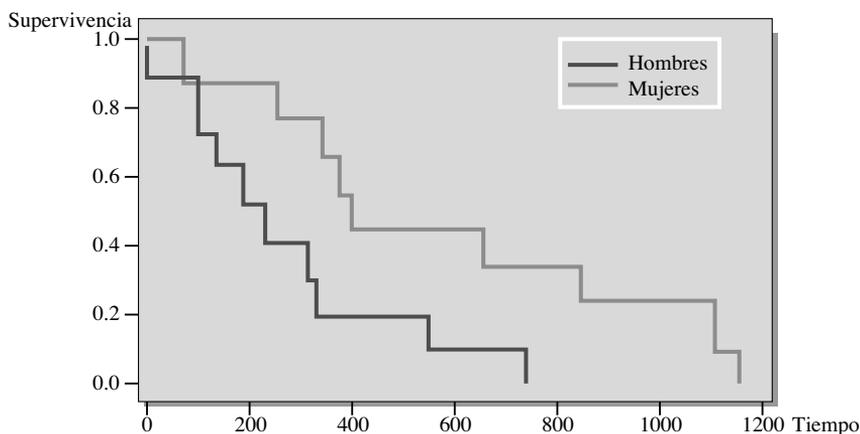


Fig.II.29.- Curvas de supervivencia de una cohorte de hombres y mujeres con cáncer.

Pero, ¿existen realmente tales diferencias o serán debidas simplemente al azar?, ¿verdaderamente el grupo de las mujeres presenta una supervivencia más alta que la de los hombres?. Nuestro interés está en ver si la aparente diferencia entre las dos curvas es estadísticamente significativa o si por el contrario ambos grupos tienen una distribución de supervivencia parecida; se trata pues de realizar un contraste de hipótesis donde la hipótesis nula es que la supervivencia de los dos grupos son iguales, aceptándola o rechazándola en función de los resultados obtenidos.

Para realizar este contraste construimos una tabla donde, para cada tiempo de muerte, esté recogido el número de fallecimientos observados en los grupos de hombres y mujeres, dados por O_H y O_M respectivamente, y el número de pacientes que continúan en observación dentro de cada grupo, es decir, el número de individuos de cada grupo que viven justo antes de cada tiempo de supervivencia; este conjunto de pacientes a riesgo en cada grupo estará representado por R_H y R_M respectivamente. Suponiendo la hipótesis nula cierta, vamos a comparar las muertes observadas con las esperadas mediante un razonamiento similar al de las tablas de contingencia y ver cómo de parecidos son estos valores. Si están muy próximos podremos aceptar como válida la hipótesis inicial y concluir que no hay diferencias entre los grupos estudiados; si por el contrario las muertes esperadas son muy diferentes a las observadas rechazaremos nuestra hipótesis nula.

Calculemos entonces el número de muertes esperadas en cada instante de tiempo en que se ha observado al menos un fallecimiento suponiendo que no existen diferencias entre el grupo de hombres y mujeres. Bajo esta hipótesis, la probabilidad de que un individuo muera en $t_1=8$ estará dada por el número total de muertes observadas en ese tiempo, ya sea hombre o mujer, dividido entre el número total de personas con vida en ese instante, es decir, $1/20$; ahora, el número de muertes esperadas en el grupo de hombres en $t_1=8$ será el producto de esta probabilidad por el número de hombres a riesgo de morir en ese instante, es decir, $11/20=0.55$ y el número esperado de muertes en las mujeres estará dado igualmente por $9/20=0.45$ puesto que 9 son las personas a riesgo dentro de este grupo.

Para el siguiente instante de tiempo, $t_1=82$, continúan en observación 18 personas, 9 hombres y 9 mujeres, puesto que un hombre murió en $t_1=8$ y otro se perdió del seguimiento a los 78 días; uno de éstos muere en $t_1=82$, luego la probabilidad de morir en este instante de tiempo es $1/18$ y el número esperado de muertes es de $9/18=0.5$ hombres y $9/18=0.5$ mujeres. De esta forma podemos ir calculando el número esperado de muertes en cada instante de tiempo que está determinado, en el caso de los hombres, por la expresión

$$\frac{O_H + O_M}{R_H + R_M} R_H$$

y en el grupo de las mujeres por

$$\frac{O_H + O_M}{R_H + R_M} R_M$$

donde O_H+O_M es el número total de muertes observadas en dicho instante de tiempo, R_H+R_M es el número total de personas con vida justo antes de ese tiempo. El número esperado de muertes en cada grupo no es más que la suma de las muertes esperadas en cada tiempo de muerte.

La Tabla II.25. muestra el proceso completo para calcular las muertes esperadas en cada grupo de pacientes, obteniéndose los valores 6.089 para los hombres y 12.91 para las mujeres.

Tabla II.25.- Cálculo de las muertes esperadas en dos grupos diferentes de pacientes con cáncer.

Tiempos de supervivencia	O_H	O_M	$O_H + O_M$	R_H	R_M	$R_H + R_M$	$\frac{O_H + O_M}{R_H + R_M} R_H$	$\frac{O_H + O_M}{R_H + R_M} R_M$
8	1	0	1	11	9	20	0.550	0.450
82	0	1	1	9	9	18	0.500	0.500
109	2	0	2	9	8	17	1.059	0.941
137	1	0	1	7	8	15	0.467	0.533
196	1	0	1	6	8	14	0.428	0.571
253	1	1	2	5	8	13	0.769	1.231
325	1	0	1	4	7	11	0.364	0.636
342	0	1	1	3	7	10	0.300	0.700
345	1	0	1	3	6	9	0.333	0.667
381	0	1	1	2	6	8	0.250	0.750
404	0	1	1	2	5	7	0.286	0.714
560	1	0	1	2	4	6	0.333	0.667
662	0	1	1	1	4	5	0.200	0.800
730	1	0	1	1	3	4	0.250	0.750
838	0	1	1	0	3	3	0.000	1.000
1114	0	1	1	0	2	2	0.000	1.000
1159	0	1	1	0	1	1	0.000	1.000
Total	10	9					6.089	12.910

Si comparamos las sumas de las muertes observadas con las esperadas en cada grupo mediante el estadístico de Pearson, como hacíamos en las tablas de contingencia, obtendremos el valor

$$T = \frac{(10 - 6.089)^2}{6.089} + \frac{(9 - 12.91)^2}{12.91} = 3.696$$

que será comparado con el valor de una distribución χ^2 con 1 grado de libertad, obteniéndose un valor P de 0.055. En consecuencia, con un error P muy próximo a 0.05, podemos rechazar la hipótesis nula de igualdad de supervivencia en los dos grupos, es decir, la supervivencia en hombres es distinta a la de las mujeres.

El test así construido recibe el nombre de Log-Rank o Mantel-Haenszel para la comparación de curvas de supervivencia y puede ser extendido sin dificultad al caso de r grupos, donde la hipótesis nula mantiene que la supervivencia es la misma en todos los grupos; el test se lleva a cabo de forma análoga, calculando las muertes esperadas en cada grupo y comparándolas con las observadas correspondientes mediante el estadístico T que en este caso tendrá r sumandos, tantos como grupos se comparen. Supuesto que la supervivencia es igual en los r grupos, el estadístico T sigue aproximadamente una χ^2 con $r-1$ grados de libertad.

No hay que perder de vista sin embargo que el tamaño de la muestra estudiada es muy importante para que el test efectuado sea «fiable» en el sentido de tener una potencia aceptable, por lo que antes de realizar cualquier comparación entre grupos es necesario asegurarnos de que el diseño de la investigación es el correcto y el tamaño muestral el adecuado; en caso contrario las conclusiones obtenidas podrían ser erróneas y en absoluto fiables. Machin²⁹ presenta tablas para el número de pacientes necesarios para realizar el test Log-Rank en ensayos clínicos, dando tablas que facilitan este trabajo.

E. ANALISIS MULTIVARIANTE

E. Sánchez-Cantalejo Ramírez

Ricardo Ocaña Riola

E. Perea-Milla López

1. Confusión e interacción

Como puede imaginar el lector, un análisis tan sencillo como el expuesto en II.C.5 acerca de la asociación entre dos variables puede presentar graves inconvenientes; pudiera ser que el Flurbiprofen fuese un tratamiento más eficaz que la combinación de Dipirona+Hioscina y que el hecho de no haber encontrado asociación se deba a que, bien por un mal diseño del estudio o bien por azar, en el grupo del Flurbiprofen hubiesen caído los enfermos con un cuadro clínico más grave que en el grupo tratado con Dipirona+Hioscina, lo que constituiría un sesgo en el diseño del estudio. Para poder comparar la eficacia de dos tratamientos deberemos estar seguros de que los dos grupos de enfermos a los que se les aplican tales tratamientos son similares en términos de cualquier variable que pueda afectar al resultado, pues en caso contrario estamos midiendo algo distinto al efecto del tratamiento: el efecto del tratamiento más el de las variables en las que los dos grupos no son similares. Si los grupos no son parecidos en cuanto a las variables que puedan afectar al resultado, no nos va a ser posible diferenciar el efecto de éstas y el del tratamiento, por lo que no podremos alcanzar el objetivo para el que se planteó el estudio.

1. 1. Confusión

Consideremos el siguiente ejemplo hipotético acerca del efecto de dos tratamientos A y B en individuos con una determinada enfermedad. Imaginemos que hacemos un seguimiento de tales enfermos y al cabo de cierto tiempo evaluamos cuantos individuos curaron con cada uno de los dos tratamientos; la Tabla II.26 muestra los resultados del estudio.

Tabla II.26.- Resultados de un estudio hipotético.

	Curación		Total
	Sí	No	
Tratamiento A	40	60	100
Tratamiento B	60	40	100
Total	100	100	200

Midiendo la asociación entre la curación y el tratamiento mediante el riesgo relativo, el cociente $(40/100)/(60/100)=0.67$ indica que es 0.67 veces más probable curar con A que con B; es decir, es $1/0.67=1.5$ veces más probable curar con el tratamiento B que con el A; realizando el contraste de hipótesis correspondiente podemos afirmar, salvo un error inferior a 0.01, que el tratamiento B cura más que el tratamiento A. Por tanto, en principio, podemos afirmar que es un 50% más probable curar con el tratamiento B que con el A, con un error $P<0.01$.

Discutiendo los resultados con los supuestos investigadores, nos dicen que el tratamiento A fué aplicado mayoritariamente en hombres, mientras la mayoría de los tratados con B fueron mujeres. Inmediatamente se plantea la siguiente pregunta: ¿qué influencia puede tener el hecho de que en los dos grupos de enfermos no haya la misma proporción de hombres que de mujeres? Para responder a esta pregunta lo que vamos a hacer es desagregar la Tabla II.26, formando dos tablas, una para hombres y otra para mujeres, como aparece en la Tabla II.27.

Tabla II.27.- Resultados de un estudio hipotético (mismos datos de Tabla II.26) desagregados por sexo.

	Hombres		Mujeres		Total
	Curado	No curado	Curado	No curado	
Tratamiento A	16	57	24	3	100
Tratamiento B	2	10	58	30	100

En esta tabla se puede apreciar que el tratamiento A fue administrado a $16+57=73$ hombres mientras el B lo fué solo a 12; si medimos ahora el riesgo relativo sólo en el grupo de los hombres resulta ser $(16/73)/(2/12)=1.32$, mientras en el grupo de muje-

res es $(24/27)/(58/88)=1.35$. En definitiva, tanto en los hombres como en las mujeres es, aproximadamente, 1.3 veces más probable curar con el tratamiento A que con el B, cosa que no ocurría en el análisis anterior en que no tenía en cuenta el sexo de los pacientes. Si no hay más variables distribuidas desigualmente en los dos grupos de tratamiento, podemos asegurar que el tratamiento A es mejor que el B tanto en hombres como en mujeres, resultado que entra en contradicción con el anterior que no tenía en cuenta el sexo de los pacientes y que mostraba una superioridad del tratamiento B sobre el A, con un riesgo relativo de 0.67.

En una situación como ésta en que el grado de asociación entre dos variables, tratamiento y curación, es distinto si se tiene o no en cuenta otra variable, el sexo, se dice que esta última es una **variable confusora** para la asociación entre el tipo de tratamiento y la curación. Si no se tiene en cuenta el sexo, o dicho de otra forma, si no se **controla** por el sexo, se dice que existe **confusión** entre las dos variables, tratamiento y curación. Lo que se acaba de hacer es estudiar la relación entre dos variables en los distintos estratos de una tercera, el sexo; a este proceso se denomina **estratificación**; evidentemente, la estratificación se debe hacer en base a más de una variable, a tantas como variables estén desigualmente distribuidas en los grupos de tratamiento y que puedan afectar al resultado. Obsérvese por último que los riesgos en los dos estratos, los grupos formados según la variable por la que se controla, son similares entre sí, 1.32 y 1.35, pero distintos al obtenido sin estratificar; ya que los riesgos en los estratos son similares, tiene sentido hablar de un único riesgo relativo para los hombres y para las mujeres.

En este ejemplo hemos controlado el efecto confusor del sexo analizando la asociación entre el tratamiento y la curación en ambos grupos de sexo; por eso se dice que en este caso se ha controlado la confusión en la fase de análisis del estudio. Otra forma de controlar el efecto confusor de las variables es en la fase de diseño, obligando en esta fase del estudio a que las variables confusoras, en nuestro ejemplo el sexo de los pacientes, se distribuya de igual forma en los dos grupos de tratamiento, es decir, diseñando un estudio apareado.

1. 2. Interacción

En la Tabla II.28. aparecen los resultados de un estudio de seguimiento de pacientes diagnosticados de melanoma cutáneo para evaluar la asociación entre supervivencia a tal cáncer y el sexo de los enfermos. Los casos se diagnosticaron entre los años 1985 y 1989; como fecha de inicio se consideró la del diagnóstico del tumor y la de finalización el 31 de diciembre de 1992. Los datos proceden del Registro de Cáncer de Granada, registro de carácter poblacional ubicado en la Escuela Andaluza de Salud Pública.

Tabla II.28.- Supervivencia de melanoma.

	Supervivencia		Total
	Sí	No	
Hombre	17	49	66
Mujer	25	31	56
Total	42	80	122

Según la Tabla II.28., una estimación del riesgo de muerte entre los hombres viene dada por el cociente $49/66$ y, para las mujeres, $31/56$; por tanto el cociente $(49/66)/(31/56)=1.34$ es el riesgo relativo de muerte de los hombres respecto a las mujeres; es decir, un hombre con melanoma está a un 34% de más riesgo de muerte que una mujer. Entre las variables que se analizaron en el estudio estaba la edad de los pacientes al diagnóstico; si disponemos de esta variable, sería deseable estudiar si este riesgo de 1.34 se mantiene, independientemente de la edad a la que se diagnosticó el melanoma en el paciente, es decir, ¿el riesgo relativo de muerte de los hombres respecto a las mujeres es el mismo sea cual sea la edad? Para responder a esta pregunta, vamos a considerar dos grupos de pacientes en función de su edad: los que tienen 45 años o menos y los que tienen más de 45 años. La Tabla II.29 presenta estos resultados. Según ésta, entre los 32 pacientes más jóvenes, los que tienen menos de 46 años, el riesgo de muerte entre los hombres es de $19/(3+19)$ y $4/(6+4)$ entre las mujeres; es decir, para este grupo de edad, el riesgo relativo de muerte de los hombres respecto a las mujeres es $(19/22)/(4/10)=2.16$. Para el otro grupo de edad, los 90 pacientes de más de 45 años, tal riesgo relativo viene dado por $(30/44)/(27/46)=1.16$; recuérdese que este riesgo era 1.34 si no considerábamos la edad de los pacientes.

Tabla II.29.- Supervivencia de melanoma según edad y sexo.

	Edad \leq 45 años		Edad \geq 45 años		Total
	Supervivencia		Supervivencia		
	Sí	No	Sí	No	
Hombre	3	19	14	30	66
Mujer	6	4	19	27	56

A la vista de estos resultados, podemos afirmar que el riesgo de muerte, independientemente de su significación estadística, es distinto en los dos estratos formados en base a la edad de los pacientes. El riesgo relativo de muerte de los hombres respecto a las mujeres depende de una tercera variable, la edad de ellos. En una situación como ésta se dice que existe **interacción** entre el sexo y la edad de los pacientes; ya que es muy distinto el riesgo de ser hombre según la edad que se tenga, no tiene sentido hablar del riesgo de ser hombre en términos absolutos, es necesario especificar la edad; habrá que dar una estimación del riesgo en el grupo de edad < 45 años y otra estimación para el riesgo en el grupo de más de 45 años.

Resumiendo, en este apartado se ha tratado con dos conceptos muy importantes en la investigación clínica. En el primer ejemplo, se ha visto que si no se tenía en cuenta el sexo de los pacientes la relación calculada, riesgo relativo 0.67, entre el tratamiento y la curación no era la real, pues tanto en los hombres como en la mujeres dicha asociación venía dada por los riesgos 1.32 y 1.35, respectivamente; ya que tal relación es muy parecida en los dos grupos, podemos decir que la asociación entre el tratamiento y la curación de la supuesta enfermedad viene dada por un valor comprendido entre 1.32 y 1.35. En definitiva, es más probable que un paciente cure con el tratamiento A que con el B, tanto en hombres como en mujeres, en contra de lo que decía el análisis que no tenía en cuenta, que no controlaba por, el sexo. Ya que en ambos grupos o **estratos** los riesgos son muy parecidos, tiene sentido combinarlos, lo que se puede conseguir con el llamado estadístico de Mantel-Haenszel³⁰.

En el segundo ejemplo, las cosas son distintas. En este caso las asociaciones entre la supervivencia al melanoma y el sexo en los dos grupos que hemos formado en función de la edad, son diferentes; por tanto, aquí no tiene sentido hablar de asociación entre supervivencia y sexo, si no se especifica a qué grupo de edad pertenece el paciente. En caso de presencia de interacción no tiene sentido dar un valor resumen de los distintos estratos pues en éstos las cosas son muy diferentes.

2. Necesidad del análisis multivariante

En el apartado anterior se ha visto cómo la relación entre dos variables puede estar condicionada por una tercera variable; si no se controla por ésta, la asociación encontrada entre las dos primeras puede ser errónea. Esta dificultad se puede obviar estudiando la relación de interés en distintos estratos de la variable por la que se quiere controlar; en el apartado II.E.1.1. sobre confusión controlamos por el sexo de los pacientes y en II.E.1.2, el apartado sobre interacción, la variable de control fué la edad. Sin embargo, en la práctica, la asociación entre dos variables puede estar influida por dos o más variables: el sexo, la edad, el estadio de la enfermedad al diagnóstico, etc.

Siguiendo con la estrategia antes expuesta, tendríamos que estudiar la relación de interés en los distintos grupos de enfermos formados en base a su sexo, edad, estadio al diagnóstico etc. El inconveniente que presenta tal estrategia es la multitud de grupos que se pueden formar; por ejemplo, imaginemos estas tres variables de las que formamos 2 grupos para el sexo, 4 para la edad y otros 4 para el estadio de la enfermedad; en definitiva serían $2 \times 4 \times 4 = 32$ grupos, por lo que la asociación de interés habrá que estudiarla en estos 32 grupos. Como quiera que en la investigación clínica el número de sujetos en estudio suele ser de moderado a bajo (en el ensayo clínico sobre el efecto del Flurbiprofen eran 52 pacientes, en el ejemplo de la supervivencia al melanoma cutáneo eran 122, etc.), ello significa que si hay muchos estratos en los que estudiar la relación de interés, el número de pacientes en cada uno de los estratos va a ser muy pequeño; así, si se quisiera controlar por sexo, edad y estadio en el ensayo del Flurbiprofen, el número medio de pacientes por estrato sería $52/32 = 1.62$. Es decir, en la mayoría de los estratos tendríamos uno o dos enfermos; el lector puede imaginar qué se puede deducir de cada uno de esos estratos con tan poca información.

Otro inconveniente de la formación de los estratos aparece que el ejemplo sobre el melanoma; hemos formado grupos respecto a los valores de la edad de los pacientes: uno constituido por los pacientes con edades ≤ 45 años, y el otro grupo lo forman los que tienen más de 45 años. Con esta formación de estratos estamos asumiendo implícitamente que la relación entre la supervivencia y el sexo es la misma en todos los pacientes con una edad inferior a 46 años, por ejemplo, en un paciente con 27 años que en uno con 44; sin embargo, ponemos en dos grupos distintos a este último paciente y a otro con 46 años; realmente esto es difícil de justificar. En definitiva, siempre que se disponga de una variable cuantitativa continua, tal como la edad, se debe tratar como tal y no categorizarla, que de alguna manera es bajarla de categoría.

El análisis multivariante es un conjunto de técnicas que permiten el análisis de los datos de una manera más eficiente que la que se acaba de exponer, en el sentido de que resuelve el problema que presentaba la estrategia de estratificación que antes se comentó. Por tanto permite, de un modo más económico, en el sentido de necesitar menos sujetos en el estudio, analizar relaciones entre dos variables controlando por tantas variables como se desee. Además permite tratar a las variables continuas como tales, no exigiendo su categorización como se tuvo que hacer anteriormente con la edad.

Por **análisis multivariante** se puede entender el análisis estadístico que trata con tres o más variables simultáneamente; bajo este criterio, la estratificación es una técnica multivariante con las deficiencias que se han comentado anteriormente. De las muchas técnicas del análisis multivariante, nosotros vamos a hablar exclusivamente de técnicas de regresión, pues el clínico suele estar interesado en estudiar cómo afec-

tan a una variable, la variable resultado, los cambios en otras variables, las predictoras; por ejemplo ¿cómo cambia el peso del recién nacido con la edad y el peso de la madre?; ¿cómo cambia el riesgo de infarto de miocardio con el cambio de los niveles de triglicéridos y de HDL; ¿cómo cambia la supervivencia del cáncer de mama según el estadio al diagnóstico y la edad de la mujer? Las técnicas de regresión multivariantes que se exponen a continuación son, con mucho, las más utilizadas en la investigación sanitaria.

Estos son tres ejemplos que se pueden analizar con las técnicas de regresión que vamos a presentar en este capítulo; la elección de estas tres se debe principalmente a la gran eclosión que se ha producido en su empleo en los últimos años en las publicaciones médicas. Algunas de ellas, como la regresión lineal múltiple, son técnicas muy antiguas, y las otras, como la regresión logística y la regresión de riesgos proporcionales, son más recientes. La razón de su gran popularidad, y no sólo en la literatura sanitaria, es debido a la disponibilidad de los ordenadores; en general, las técnicas multivariantes implican cálculos muy complicados que pueden llegar a implicar millones de operaciones aritméticas lo que, evidentemente, sólo una máquina es capaz de llevar a cabo.

La elección de una determinada técnica de regresión va a venir condicionada por la naturaleza de la variable respuesta. Si ésta es continua, por ejemplo, el peso del recién nacido, el modelo a utilizar será el de regresión lineal; cuando la respuesta sea dicotómica, presencia o no de infarto, no se pueden utilizar los métodos diseñados para variables continuas, y se utilizará el modelo de regresión logística; por último, cuando la respuesta sea el tiempo que transcurre hasta la presentación de un determinado suceso en los pacientes, por ejemplo la muerte del sujeto, se debe utilizar alguno de los llamados modelos de supervivencia, de los cuales el de riesgos proporcionales es el más empleado en la literatura médica. Como elemento común a las tres técnicas está lo que podemos llamar el componente sistemático o combinación de las predictoras

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

donde X_1, X_2, \dots, X_p son las p variables predictoras y $\beta_0, \beta_1, \dots, \beta_p$ son parámetros a estimar y que nos darán idea de la asociación de las predictoras con la variable resultado; la interpretación de estos parámetros dependerá del modelo elegido, como se verá más adelante. En cuanto a las predictoras, no hay ninguna restricción acerca de su naturaleza, pudiendo ser categóricas o numéricas, tanto discretas como continuas en cualquiera de los tres modelos.

3. Regresión lineal múltiple

En el apartado II.D.3 se presentó el modelo de regresión lineal simple como método para el estudio de la relación entre dos variables cuantitativas: los niveles de hemoglobina y creatinina; el objetivo era estudiar cómo cambia la variable resultado, la hemoglobina, con el cambio de la predictor, la creatinina. Este mismo objetivo se puede enunciar de esta otra forma: ¿qué parte de la variabilidad del nivel de hemoglobina en los individuos viene explicada por los distintos niveles de creatinina?; en II.D.3. se vió que el coeficiente de determinación era 0.71, lo que se interpretaba como que el 71% de la variabilidad de los niveles de hemoglobina venía explicada por los distintos niveles de creatinina. Por tanto, queda un 29% de la variabilidad de la hemoglobina que no es achacable a los distintos niveles de creatinina, por lo que es de interés encontrar otras variables distintas a la creatinina que puedan ayudar a explicar ese 29% restante.

El modelo de regresión lineal múltiple permite estudiar la relación entre una variable resultado continua y un conjunto de dos o más predictoras, pudiendo ser éstas de cualquier naturaleza; si notamos por Y la variable resultado y por X_1, X_2, \dots, X_p las predictoras, tal modelo establece que la media de la variable resultado $E(Y)$ se puede expresar por la relación

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

donde los β_i son los parámetros del modelo, a estimar a partir de la muestra de sujetos elegida; cuando se dispone de una sola predictor este modelo se reduce al de la regresión lineal simple. Recuérdese que en este caso los parámetros se estimaban por el método de los mínimos cuadrados; en el modelo multivariante es el mismo método de estimación pero más complicado, aunque sólo desde el punto de vista del cálculo, cuestión que no debe preocupar, pues eso lo realiza el ordenador.

La diabetes mellitus es una enfermedad de alta prevalencia en los países desarrollados que puede llevar a complicaciones clínicas graves a los que la padecen; existen evidencias que un buen control de la enfermedad reduce el riesgo de tales complicaciones. Pringle et al ³¹ presentan un estudio para evaluar posibles factores asociados al control de la diabetes que tienen que ver con las características del paciente, del médico y del proceso de atención; entre ellas podemos citar: 1) tipo de tratamiento: insulina, antidiabéticos orales o dieta; 2) sexo del paciente; 3) años desde el diagnóstico de la enfermedad ($>7, <7$); 4) número de enfermedades relacionadas con la diabetes ($>1, 1$); 5) nivel socioeconómico; 6) estilos de vida (tabaco, alcohol); 7)

edad y sexo del médico; 8) actitud y conocimiento de la diabetes por parte del médico; 9) existencia de protocolo para el tratamiento de la enfermedad, etc. Como medida del control de la enfermedad utilizan el porcentaje de hemoglobina glicosilada (hemoglobina A_{1c}) de tal manera que a mayor valor de esta variable peor control existe de la enfermedad.

3. 1. Tratamiento de las predictoras categóricas

Como se ha dicho anteriormente, no existe ningún tipo de restricción acerca de la naturaleza de las predictoras de un modelo de regresión multivariante; de todas formas cuando se dispone de variables predictoras categóricas, éstas no pueden ser tratadas tal cual y es necesario un tratamiento especial como se expone a continuación. Como ocurre con el resto de variables, a la hora de construir los ficheros para poder ser tratados informáticamente, es necesario asignar valores numéricos a las distintas categorías, pero éstos no son más que etiquetas para distinguir un tipo de tratamiento de otro. Consideremos en primer lugar una variable dicotómica, por ejemplo, el sexo de los pacientes; para este tipo de variables a una de las categorías se le suele asignar el valor 1 y a la otra el valor 0; aunque otras asignaciones de valores son igualmente lícitas, ésta es la más común pues es la que hace más fácilmente interpretables los coeficientes del modelo como más tarde se verá.

La variable tipo de tratamiento necesita un comentario especial; ésta es una variable categórica con tres niveles: insulina, antidiabéticos orales y dieta. En el estudio se le asignó el valor 1 a los tratados con insulina, a los tratados con antidiabéticos orales se les asignó el 2 y a los que sólo se trataron con dieta se les asignó el valor 3; otra posible asignación, tan válida como la anterior, podría haber sido 1, 7, y 98 ya que, como se acaba de decir, no son más que etiquetas, no tienen sentido estrictamente numérico. Como quiera que cuando el ordenador encuentra un número lo trata como tal, es necesario algún artificio para tratar las variables categóricas en los modelos de regresión. Tal artificio consiste en crear lo que se denominan **variables indicadoras (dummy)**; para el caso que nos ocupa es necesario crear 2 variables indicadoras como sigue: una primera variable que toma el valor 1 si el paciente se trató con antidiabéticos orales y 0 en caso contrario; una segunda variable que vale 1 si el tratamiento consistió solamente en la dieta y 0 en caso contrario. Con estas dos variables indicadoras se puede identificar el tratamiento de cualquier paciente; así, un paciente identificado por (0,0) no habrá sido tratado ni con antidiabéticos orales ni con dieta, es decir, habrá sido tratado con insulina; (1,0) identificará a un paciente tratado con antidiabéticos orales y (0,1) a uno tratado sólo con dieta. La categoría notada con (0,0) es la categoría de referencia. En general, cuando se disponga de una predictoras categórica con k categorías, será necesaria la creación de $k-1$ variables indicadoras.

3. 2. Interpretación de los coeficientes del modelo

En base a una muestra de 308 pacientes, se estima la asociación entre la hemoglobina A_1 y el conjunto de las predictoras mediante un modelo de regresión lineal múltiple pues la variable resultado, la hemoglobina A_1 , es una variable continua. Los autores presentan los resultados de la Tabla II.30.; en ella aparece una primera columna en donde están las variables predictoras relacionadas estadísticamente con la variable resultado, otra columna donde están las estimaciones de los coeficientes correspondientes a las variables anteriores y, por último, una columna con los errores estándar de tales estimaciones. Aunque algunas de las variables de la primera columna son numéricas (años desde el diagnóstico y número de enfermedades relacionadas con la diabetes) los autores las consideraron como dicotómicas según los criterios que se acaban de señalar.

Para poder interpretar los resultados es fundamental conocer cómo se han asignado los valores a las distintas categorías de las variables; en la notación que utilizan los autores del trabajo para la variable sexo (mujer/hombre) significa que a las mujeres le han asignado valor 1 y a los hombres el 0; a los que han tenido acceso al dietista hospitalario le asignaron valor 1 y a los que no tuvieron acceso valor 0; etc.

Tabla II.30.- Variables asociadas al nivel del hemoglobina A_1 .

Variable	Estimación	Error estándar
Constante	10.51	0.59
Antidiabéticos orales/insulina	0.29	0.37
Dieta/insulina	-0.97	0.46
Nº enfermedades relacionadas con la diabetes ($>1/\leq 1$)	0.20	0.08
Sexo (mujer/hombre)	0.72	0.28
Años desde el diagnóstico ($>7/\leq 7$)	0.04	0.02
Acceso al dietista hospitalario (sí/no)	-1.06	0.43
Médico interesado en la diabetes (sí/no)	-0.86	0.38

A diferencia del modelo simple donde β_1 medía el cambio que experimenta la media de la variable resultado Y por unidad de aumento de la predictora X_1 , en el modelo multivariante β_i mide el cambio que experimenta la media de la variable resultado por unidad de aumento de la predictora X_i , a igualdad de valores de las restantes variables presentes en el modelo.

Según la Tabla II.30., el modelo de regresión estimado se puede escribir así

$$E(HbA_1) = 10.51 + 0.29(\text{Ant. orales/insulina}) - 0.97(\text{Dieta/insulina}) + 0.20(\text{N}^\circ \text{enferm.}) \\ + 0.72(\text{Sexo}) + 0.04(\text{Años diagnost.}) - 1.06(\text{Acc.dietit.}) - 0.86(\text{Interés})$$

donde $E(HbA_1)$ significa la media de hemoglobina A_1 . Como se puede observar, para la variable tratamiento que era categórica con tres categorías, aparecen dos variables indicadoras, como se explicó en el apartado anterior.

Veamos entonces cómo interpretar los resultados de este estudio. En primer lugar, toda variable que tenga un coeficiente estimado con signo positivo significa que a mayor valor de tal predictora mayor valor de la variable respuesta. El valor 0.72 correspondiente al sexo se interpreta de la siguiente forma: a igualdad de tipo de tratamiento, de número de enfermedades relacionadas con la diabetes, años desde el diagnóstico, acceso al distista hospitalario e interés del médico en la diabetes, una mujer (valor 1 de la variable sexo) tiene por término medio 0.72 unidades de hemoglobina A_1 más que un hombre (valor 0 de la variable sexo); es decir, a igualdad en las restantes variables que aparecen en el modelo, las mujeres tienen peor control, mayor porcentaje de hemoglobina glicosilada, que los hombres. El valor 0.72 estima el cambio medio en hemoglobina A_1 de las mujeres respecto de los hombres, la categoría de referencia.

La estimación 0.29 significa que a igualdad de las restantes variables, los pacientes tratados con antidiabéticos orales tienen 0.29 unidades más de hemoglobina A_1 que los tratados con insulina; de igual forma, los tratados sólo con dieta tienen por término medio 0.97 unidades menos que los tratados con insulina. Aquí se puede observar que se comparan tanto los tratados con antidiabéticos orales como los tratados con dieta, respecto a los tratados con insulina; recuérdese que los tratados con insulina eran los sujetos identificados con la pareja (0,0) que se eligió como categoría de referencia.

El valor -0.86 correspondiente a la variable que mide el interés del médico por la diabetes se puede interpretar como que a igualdad de valores en las restantes variables, las personas tratadas por un médico interesado en la diabetes tienen por término medio 0.86 unidades menos de hemoglobina A_1 que las personas tratadas por un médico sin especial interés en esta enfermedad; es decir, el hecho de que el médico tenga un especial interés en la diabetes está asociado a un mejor control de la enfermedad en sus pacientes.

Hasta ahora hemos interpretado el efecto de una sola variable, es decir, hemos comparado mujeres con hombres, tratados por un médico con interés en la diabetes frente a los tratados por un médico sin especial interés en esta enfermedad, etc. Tam-

bién es posible la comparación entre dos diabéticos que difieran en dos o más variables de las que aparecen en la Tabla II.30.; por ejemplo, ¿qué diferencia promedio existe entre un hombre (sexo=0), tratado con insulina (tratamiento=(0,0)), con tres enfermedades relacionadas con la diabetes (Nº enfermedades=1), con una diabetes diagnosticada hace 8 años (Años desde el diagnóstico=1), con acceso al dietista (Acceso=1) y tratado por un médico interesado en la diabetes (Interés=1), respecto a una mujer (sexo=1), tratado con dieta (tratamiento=(0,1)), con 1 enfermedad relacionada con la diabetes (Nº enfermedades=0), con una diabetes diagnosticada hace 3 años (Años desde el diagnóstico=0), sin acceso al dietista (Acceso=0) y tratada por un médico interesado en la diabetes (Interés=1)? Según el modelo estimado, el nivel medio de hemoglobina A₁ para los hombres de las características del anteriormente descrito es

$$E (HbA_1)_{Hombre} = 10.51 + 0.29(0) - 0.97(0) + 0.20(1) + 0.72(0) + 0.04(1) - 1.06(1) - 0.86(1)$$

y el nivel medio para una mujer con las características de la anterior es

$$E (HbA_1)_{Mujer} = 10.51 + 0.29(0) - 0.97(1) + 0.20(0) + 0.72(1) + 0.04(0) - 1.06(0) - 0.86(1)$$

por lo que la diferencia entre ambos valores es $-0.97(0-1)+0.20(1-0)+0.72(0-1)-1.06(1-0)=-0.61$; es decir, los hombres del tipo anterior tienen por término medio 0.61 unidades de hemoglobina A₁ menos que las mujeres con las características antes descritas.

En tercera columna de la Tabla II.30. aparecen los errores estándar de las estimaciones de los parámetros del modelo; como todo error estándar, éstos dan idea de la fiabilidad de las estimaciones y nos permiten tanto evaluar la significación estadística de la asociación entre las predictoras con la variable resultado como la construcción de intervalos de confianza para los parámetros del modelos. Así, dividiendo cada estimación por su error estándar y comparando tal cociente en la distribución normal nos permite evaluar el error P; por ejemplo, el cociente entre la estimación 0.72 correspondiente al sexo y su error estándar 0.28, $0.72/0.28=2.57$ nos permite afirmar que el sexo está relacionado con el control de la diabetes con un error aproximadamente igual a 0.01. Obsérvese como el cociente $0.29/0.37=0.78$ no es significativo ni a un error tan alto como el 0.40, por lo que no hay evidencias para afirmar que existan diferencias, en términos del control de su diabetes, entre los individuos tratados con antidiabéticos orales y los tratados con insulina.

También podemos construir intervalos de confianza para cada uno de los parámetros correspondientes a las variables presentes en el modelo. Como se dijo en el

apartado II.C.3, los intervalos de confianza para un parámetro cuyo estimador siga una distribución normal tiene como expresión general

$$\text{estimador} \pm z(e.e)$$

donde z es el valor buscado en la distribución normal al nivel de confianza elegido y $(e.e.)$ es el error estándar del estimador. Así, $-0.86 \pm 1.96 (0.38) = (-1.60, -0.12)$ es un intervalo, al 95% de confianza, para el efecto medio del interés del médico en la diabetes; esto se puede interpretar diciendo que con una confianza del 95%, el hecho de que el médico esté especialmente interesado en la diabetes, hace que sus enfermos tengan por término medio como poco 0.12 y como mucho 1.6 unidades menos de hemoglobina A_1 que los enfermos atendidos por médicos no interesados por esta enfermedad, controlando por el efecto de las restantes variables presentes en el modelo.

3.3. Selección del mejor modelo

En la práctica de la investigación clínica es frecuente medir muchas variables que a priori se piensa que pueden estar relacionadas con la variable resultado; los autores del trabajo que venimos comentando midieron más de 20 variables para tratar de explicar las diferencias en hemoglobina A_1 entre los diferentes pacientes. La cuestión que ahora se plantea es decidir cuales de las variables medidas son relativamente importantes a la hora de predecir el valor de la variable resultado o, dicho de otra forma, qué variables son las que ayudan a explicar los diferentes niveles de hemoglobina A_1 en los distintos pacientes; este proceso se conoce con el nombre de **selección de las variables** predictoras; obsérvese como de entre las variables medidas en el estudio, en el modelo de regresión de la Tabla II.30. tan solo aparecen 6 variables: el tipo de tratamiento, el número de enfermedades que tienen que ver con la diabetes, el sexo del paciente, años desde el diagnóstico, acceso al dietista y el interés del médico por la diabetes.

Existen distintas estrategias a la hora de seleccionar variables; una de ellas está basada en los denominados algoritmos **paso a paso (stepwise)** muy populares en la literatura sanitaria. Uno de ellos, el llamado **hacia adelante (forward stepwise)** procede de la siguiente forma: en un primer paso se estudia la relación entre la variable resultado y cada una de las predictoras sin tener en cuenta las restantes variables medidas, es decir, se realizan análisis de regresión simple, tantas como predictoras se hayan medido y se elige aquella predictora que muestre más asociación con la variable resultado, es decir, aquella que dé lugar a un coeficiente de determinación más alto. En un segundo paso, de entre las restantes variables se elige aquella que añada más información sobre la que ya presta la primera variable elegida, es decir, como segunda

variable se elige aquella que explique la mayor cantidad de la variabilidad que no explicaba la primera variable elegida. Este proceso se repite hasta que no haya ninguna variable que añada información sobre la que aporten las variables anteriormente seleccionadas. En cada paso, para evaluar si la información aportada por una variable es o no importante se utiliza algún criterio estadístico de inclusión, por ejemplo, $P < 0.05$.

Un segundo algoritmo es el llamado **hacia atrás (backward stepwise)** que consiste en construir en primer lugar un modelo que contenga como predictoras a todas las variables medidas e ir quitando de él las variables que no aporten información estadísticamente significativa hasta llegar a un modelo en el que todas las variables sean estadísticamente significativas; de igual manera se debe establecer algún criterio estadístico para ir eliminando las variables. Una combinación del método hacia adelante y el hacia atrás es llamado **paso a paso (stepwise)** a secas, que quizás sea el más utilizado. De todas formas, sólo en muy contadas ocasiones se deben utilizar estos métodos automáticos de selección de variables; en la mayoría de las ocasiones el clínico está interesado en evaluar el efecto de una o más predictoras y, por tanto, éstas deben entrar en el modelo, independientemente de su significación estadística³².

Todo modelo estadístico se puede aplicar si y solo si se cumplen una serie de hipótesis acerca de las variables medidas; en caso de que éstas no se cumplan las conclusiones que se derivan del modelo construido no son válidas por lo que es necesario la comprobación de su verificación. El conjunto de técnicas relacionadas con esta cuestión se conoce como **diagnóstico** del modelo. Aunque esto es un tema que sale fuera del nivel de la presente monografía es una cuestión de gran importancia y, desgraciadamente, muchas veces olvidada; de nuevo la colaboración con un profesional de la estadística se hace imprescindible.

4. Regresión logística múltiple

Horbar et al.³³ presentan los resultados de un estudio diseñado para evaluar el riesgo de muerte en recién nacidos con pesos entre 501 y 1500 gramos. Para ello se diseñó un estudio de cohorte en el que se incluyeron 3603 niños con pesos comprendidos entre los límites anteriores; el periodo de seguimiento fué de 2 años, de 1 de Noviembre de 1987 a 31 de Octubre de 1989. Como variable resultado se consideró el fallecimiento o no del recién nacido, variable dicotómica, y como predictoras se consideraron el peso en gramos, el tamaño según la edad gestacional, la raza, el sexo, tipo de parto (vaginal/cesárea), presentación del parto, el índice Apgar al minuto, presencia o no de complicaciones de la madre (diabetes insulín-dependiente, eclampsia, hipertensión o hemorragia) antes del parto.

En II.D.4 se expuso el modelo de regresión logística simple como método para estudiar la asociación entre una variable resultado dicotómica y una sola predictora. Ahora disponemos de varias predictoras y de lo que se trata es de estudiar la relación entre la variable resultado y el conjunto de todas las predictoras simultáneamente. Los autores del estudio utilizan el modelo de **regresión logística múltiple** para tal fin. De modo similar al modelo de regresión lineal múltiple, el modelo logístico multivariante se representa como sigue

$$\text{Log} \frac{P}{1 - P} = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

ecuación similar al modelo logístico univariante con la única diferencia de que ahora aparecen varias variables predictoras y donde P es también la probabilidad de presentar la característica de interés, en este caso, la muerte del recién nacido. Una forma alternativa de expresar el mismo modelo es

$$P = \frac{e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}}$$

que, en el caso de un estudio de seguimiento, permite calcular la probabilidad de que un individuo con determinados valores de las predictoras presente la característica de interés.

Una diferencia fundamental entre el caso uni y multivariante es la que concierne a la interpretación de los parámetros del modelo. En el caso univariante el coeficiente β_i se podía interpretar como el cambio que experimenta el $\text{logit}(P)$ por unidad de aumento de la variable X_i ; ahora, en el modelo logístico multivariante, el coeficiente i es el cambio que experimenta el $\text{logit}(P)$ por unidad de aumento de la variable X_i , a igualdad de valores en las restantes variables presentes en el modelo; dicho de otra forma, controlando por las restantes variables. Una definición equivalente establece que e^{β_i} es la *O.R.* entre dos individuos que se diferencian en una unidad en términos de la variable X_i pero siendo iguales respecto a las demás variables. En general la *O.R.* entre dos individuos con valores x_{i1} y x_{i2} de la variable X_i , pero con iguales valores de las restantes variables, viene dada por la expresión

$$O.R. = e^{\beta_i(x_{i1} - x_{i2})}$$

Como siempre, los parámetros del modelo no los vamos a poder conocer exactamente, por tanto a lo que tendremos acceso es a sus estimaciones en base a los 3603 individuos del estudio. De la misma manera que en el caso univariante, el método de

estimación de los parámetros es el de máxima verosimilitud que, como ya se comentó, conlleva gran complejidad de cálculo por lo que se hace imprescindible el auxilio de un ordenador.

En principio, no todas las variables medidas en los recién nacidos tienen porqué ayudar a predecir la probabilidad de muerte; como en el modelo de regresión lineal múltiple se plantea la cuestión de seleccionar las variables predictoras que aporten información, estadísticamente significativa, a la hora de predecir la probabilidad de muerte del recién nacido. Los autores del estudio utilizaron el procedimiento de selección hacia atrás (backward), llegando al modelo estimado que aparece en la Tabla II.31.

Tabla II.31.- Estimación del modelo para predecir la mortalidad en recién nacidos con pesos entre 501 y 1500 grs.

Variable	Estimación	Error estándar
Constante	2.606	
Peso en grs./100	-0.422	0.029
BPSEG ^a	-0.491	0.157
Raza	-0.450	0.145
Sexo	0.656	0.131
Apgar	0.971	0.131

Veamos cómo interpretar estos resultados; en primer lugar es fundamental conocer la forma en que se codificaron las variables categóricas pues de tal codificación depende la interpretación de los coeficientes. La variable resultado se codificó como 1 si murió el recién nacido y 0 si sobrevivió; el peso del recién nacido se expresó en gramos/100, por lo que el coeficiente de esta variable expresará cambio en el $\text{logit}(P)$ por incremento de 100 gramos de peso; esto lo hacen los autores porque es esperable que tal cambio por gramo de aumento en el peso sea desprezable. El bajo peso según la edad gestacional, BPSEG, se codificó como 1 si el peso del recién nacido estaba por debajo del percentil 10 para su edad gestacional y 0 en caso contrario; la raza: 1 si era de raza negra y 0 en caso contrario; el sexo: 1 hombre 0 mujer y el Apgar, 1 si era ≤ 3 y 0 si era > 3 . Con estas asignaciones, el modelo estimado se puede interpretar como sigue: un recién nacido con bajo peso para su edad gestacional, BPSEG=1, respecto a otro que esté por encima del percentil 10 del peso para su edad gestacional, BPSEG=0, está a $e^{-0.491(1-0)}=0.61$ veces más riesgo, medido en términos de *O.R.*, de muerte, con la condición de que esos dos niños sean iguales en cuanto a las restantes variables presentes en el modelo: el peso, la raza, el sexo y el Apgar; como quiera que este riesgo es menor que la unidad eso significa que el recién nacido con bajo peso para su edad gestacional está a menor riesgo de muerte. Dicho de otra forma, el niño que no tiene

bajo peso para su edad gestacional está a $1/0.61=1.63$ veces más riesgo de muerte que el que tiene bajo peso para su edad gestacional, con la condición de que sean iguales en términos de las otras cuatro variables presentes en el modelo ajustado.

Dos recién nacidos iguales en peso, BPSEG, sexo y Apgar, pero uno de raza negra y el otro de otra raza, el primero está a $e^{-0.45(1-0)}=0.64$ veces más riesgo de muerte que el segundo; es decir, los niños de razas distintas a la negra están a $1/0.64=1.56$ veces más riesgo de morir que los de raza negra. El papel del peso del recién nacido sobre el riesgo de muerte es como sigue: a igualdad en los valores de las restantes variables, por cada aumento de 100 gramos, pues en el modelo ajustado no aparece el peso en gramos sino el peso en gramos dividido por 100, en el peso de un recién nacido el riesgo de muerte aumenta en $e^{-0.422}=0.66$ veces, es decir, el riesgo disminuye $1/e^{-0.422}=1.52$ veces, o sea, el riesgo decrece en un 52%. Si un recién nacido pesó 850 gramos y otro 1250 gramos, el riesgo de muerte del primero es $e^{-0.422(850/100-1250/100)} = e^{1.69} = 5.4$ veces superior al del segundo, siempre que sean iguales en BPSEG, raza, sexo y Apgar.

Merece la pena el siguiente comentario: que el niño de 850 gramos esté a 5.4 veces más riesgo de muerte que el de 1250 es válido sean los dos de bajo peso o los dos de peso apropiado, los dos de raza negra o los dos de raza distinta a la negra, los dos niños o los dos niñas y, por último, los dos con un Apgar ≤ 3 o los dos > 3 . Es decir, este riesgo no cambia en función de los valores de estas cuatro variables. Los autores no hacen ningún comentario acerca de la posible interacción entre las cinco variables por lo que no es posible saber si es que no hay interacción o es que no la estudiaron; imaginemos que el término de interacción entre el peso del recién nacido y la raza hubiese sido significativo. El tal caso ya no podríamos decir que el riesgo del que pesó 850 gramos es 5.4 veces que el riesgo del que pesó 1250; habría que hablar de un riesgo en caso de que los dos fuesen niños y de otro riesgo distinto en el caso de que fuesen niñas. Si existe interacción entre dos variables, no se puede hablar del riesgo de una de ellas pues hay que especificar el valor de la variable con la que interactúa.

Aparte de poder comparar dos individuos respecto a una variable pero siendo iguales en relación a las restantes variables presentes en el modelo ajustado, es posible la comparación de dos individuos cualesquiera. Imaginemos una niña de raza negra, con 1150 gramos al nacer, con bajo peso para su edad gestacional y con un Apgar ≤ 3 ; cuántas veces está a más riesgo de muerte que un niño que no sea negro, con 1400 gramos de peso, con peso normal para su edad gestacional y con un Apgar > 3 ? En general, el riesgo entre dos individuos viene dado por la expresión

$$O.R. = e^{\beta_1(x_{11}-x_{12}) + \beta_2(x_{21}-x_{22}) + \dots + \beta_p(x_{p1}-x_{p2})}$$

Así, para el primer individuo, la niña, el valor de la primera variable $x_{11}=11.50$ pues en el modelo se ha considerado el peso en gramos dividido por 100; el valor de la segunda variable, la BPSEG, es $x_{21}=1$, la raza es $x_{31}=1$ porque es negra, $x_{41}=0$ porque es niña y $x_{51}=1$ porque tiene un Apgar ≤ 3 . Para el segundo individuo, $x_{12}=14.00$, $x_{22}=0$ porque tiene peso por encima del percentil 10 para su edad gestacional, $x_{32}=0$ porque no es negro, $x_{42}=1$ porque es un niño y $x_{52}=0$ porque tiene un Apgar >3 . Por tanto, una estimación del riesgo de morir de la niña respecto del niño será:

$$e^{-0.422(11.5-14.0) - 0.491(1-0) + 0.656(0-1) + 0.971(1-0)} = e^{0.429} = 1.54$$

es decir, la niña tiene un 54% más riesgo de morir que el niño.

Además de poder comparar riesgo de muerte entre dos individuos cualesquiera, en los estudios de seguimiento, el modelo logístico multivariante permite calcular la probabilidad de muerte individual; esto no ocurre si el diseño empleado hubiese sido un estudio de casos y controles, pues en ese caso el coeficiente β_0 no es interpretable; como quiera que el diseño empleado en este estudio fué de cohorte, podemos calcular la probabilidad de muerte de cualquier recién nacido. Así, para la niña anterior la probabilidad de muerte es

$$P = \frac{e^{2.606 - 0.422(11.5) - 0.491(1) - 0.45(1) + 0.656(0) + 0.971(1)}}{1 + e^{2.606 - 0.422(11.5) - 0.491(1) - 0.45(1) + 0.656(0) + 0.971(1)}} = 0.098$$

lo que significa que aproximadamente el 10%, exactamente el 9.8%, de las niñas con las características de esta niña, morirán.

Supongamos otro recién nacido de 750 gramos de peso, con peso normal para su edad gestacional, BPSEG=0, de raza blanca, niño y con Apgar inferior a 4; la probabilidad de muerte para este niño es

$$P = \frac{e^{2.606 - 0.422(7.5) - 0.491(0) - 0.45(0) + 0.656(1) + 0.971(1)}}{1 + e^{2.606 - 0.422(7.5) - 0.491(0) - 0.45(0) + 0.656(1) + 0.971(1)}} = 0.74$$

lo que significa que, aproximadamente el 75%, 3 de cada 4 niños con tales características, morirán.

Evidentemente todo lo dicho anteriormente acerca de los criterios de selección de variables y diagnóstico del modelo son aplicables aquí.

Catalona et al²⁸ presentan un estudio donde aplican el modelo de regresión logística multivariante para evaluar la importancia del PSA como predictor del cáncer de próstata; después de controlar por variables como la edad del paciente, los resultados del examen rectal y de la ultrasonografía, el riesgo de padecer cáncer de próstata en un individuo con $\text{PSA} \geq 4 \mu\text{g/l}$ es, aproximadamente, seis veces superior al riesgo en un paciente con niveles inferiores a $4 \mu\text{g}$.

5. Análisis de supervivencia: el modelo de Cox

Cuando se plantea un análisis de datos donde interviene un único factor que divide la cohorte de individuos en dos o más grupos diferentes, el método de Kaplan-Meier permite construir una curva de supervivencia para cada grupo que podrá ser comparada con las restantes como vimos en II.D.6.2. Sin embargo, en análisis de mayor envergadura es frecuente disponer de varias variables o factores pronóstico que potencialmente pueden modificar la supervivencia de cada individuo en función de los valores que tomen; de esta forma si quisiéramos ver conjuntamente la influencia de la Edad y Estadío clínico (con niveles de I a IV) en la supervivencia de una cohorte de pacientes afectados de cáncer de mama, una posible estrategia basada en curvas de Kaplan-Meier sería categorizar la variable Edad considerando por ejemplo tres grupos de pacientes con edades comprendidas entre 30-49 años, 50-69 años o más de 70 años, cruzar cada uno de los grupos de pacientes obtenidos con los cuatro niveles de Estadío considerados, construir una curva de supervivencia mediante el método Kaplan-Meier para cada uno de los 12 grupos de pacientes resultantes y compararlas entre sí a través de la superposición de gráficas.

Como se ha repetido anteriormente, esta metodología no es la más eficiente por la enorme cantidad de sujetos de estudio que requiere, debido al gran número de grupos de individuos distintos que pueden presentarse. Por todo ello es necesario recurrir a modelos multivariantes más complejos que permitan investigar el efecto simultáneo de las variables explicativas o factores pronóstico en la supervivencia de los pacientes, así como comparar diferentes grupos de sujetos, siendo el modelo de Cox uno de los métodos más populares en la investigación clínica que resuelve este tipo de problemas.

Para comprender la técnica utilizada por el modelo de Cox es necesario introducir el concepto de **función o tasa de riesgo** en una cohorte de pacientes a la que llamaremos en adelante $h(t)$, siendo t el tiempo que se expresará en días, meses o años dependiendo del problema concreto que se estudie. En el apartado II.D.5 vimos que la función de supervivencia, $S(t) = P[T > t]$, representaba la probabilidad de que se pro-

duzca una muerte en un instante de tiempo posterior a un tiempo dado t , por lo que si queremos calcular la probabilidad de que un paciente permanezca con vida transcurridas t unidades de tiempo desde el inicio del seguimiento, bastará con calcular $P[T > t]$; sin embargo si quisiéramos saber qué probabilidad hay de que un paciente muera en los próximos 30 días, tendremos que especificar además cuanto tiempo ha pasado desde el inicio del seguimiento, ya que, en igualdad de condiciones, un individuo que acaba de ser diagnosticado de cáncer tendrá presumiblemente muchas más posibilidades de continuar con vida un mes después que otro con la enfermedad diagnosticada hace tiempo. Por ello no es suficiente con preguntarse qué probabilidad existe de morir transcurridos 30 días, es necesario indicar cuanto tiempo ha sobrevivido el paciente hasta el momento; en realidad lo que buscamos no es más que la probabilidad de que ocurra una muerte antes del instante $t + \Delta t$ si el paciente ha sobrevivido hasta el instante t , o dicho de otra forma, la probabilidad de morir entre los instantes de tiempo t y $t + \Delta t$, sabiendo que el paciente continuaba con vida en el tiempo t . Este resultado se expresa matemáticamente mediante la probabilidad condicionada $P[T < t + \Delta t | T > t]$ donde Δt representa el incremento de tiempo transcurrido, los 30 días.

Pero, ¿qué ocurriría si Δt fuera muy pequeño, casi despreciable? En ese caso tendremos una probabilidad de muerte instantánea puesto que el intervalo definido por t y $t + \Delta t$ es tan pequeño que casi puede considerarse como un instante de tiempo concreto. Definiremos entonces la función de riesgo $h(t)$ mediante la expresión

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P [T \leq t + \Delta t | T > t]}{\Delta t}$$

que representa el riesgo de muerte instantáneo para un paciente que ha sobrevivido hasta el tiempo t . Lógicamente entre la función de supervivencia y de riesgo existe una relación inversa, es decir, a mayor riesgo de muerte menor probabilidad de sobrevivir.

El modelo de Cox consiste en introducir factores pronóstico que expliquen la variación del riesgo de los pacientes mediante la expresión

$$h(t, X_1, \dots, X_p) = h_0(t)e^{\beta_1 X_1 + \dots + \beta_p X_p}$$

donde X_1, X_2, \dots, X_p son los factores pronóstico introducidos y $h_0(t)$, llamada **función de riesgo base**, es el valor de la función de riesgo cuando se anula el efecto de las variables, es decir, cuando $X_1, X_2, \dots, X_p = 0$; el interés del modelo está en la posibilidad de comparar los riesgos entre pacientes con características diferentes, enten-

diendo por características los distintos valores que pueden tomar las variables consideradas como factores pronóstico.

La pregunta que necesitamos responder ahora es: ¿qué interpretación dar a los coeficientes β de este modelo? Para ello empecemos con el caso más simple; un estudio en el que intervienen una única variable X predictora del riesgo de muerte, por ejemplo, presencia o no de ascitis, y supongamos que la presencia de ascitis se ha codificado con un 1 y la ausencia con un 0. Para una sola predictora, el modelo de Cox adoptaría la forma

$$h(t, X) = h_0(t)e^{\beta X}$$

El riesgo de muerte para un individuo con ascitis, $X=1$ vendrá dado por

$$h(t, 1) = h_0(t)e^{\beta(1)}$$

y, análogamente, el riesgo de muerte para un sujeto sin ascitis $X=0$, será

$$h(t, 0) = h_0(t)e^{\beta(0)}$$

por lo que, si hacemos el cociente de ambos riesgos, obtendremos lo que se denomina **riesgo relativo (hazard ratio)** dado por

$$\frac{h(t, 1)}{h(t, 0)} = \frac{h_0(t)e^{\beta(1)}}{h_0(t)e^{\beta(0)}} = e^{\beta}$$

es decir, en un modelo de Cox univariante, β es un número tal que e^{β} es el riesgo relativo de un paciente en relación a otro con una unidad menos respecto a la variable predictora. Christensen³⁴ presenta un estudio en donde analiza el efecto aislado de la cantidad de albúmina en sangre, expresada en gramos/litro, sobre la supervivencia de pacientes con enfermedad hepática grave; el coeficiente de esta predictora es -0.42; ¿cómo interpretar este resultado? Como se acaba de decir, $e^{-0.42}=0.657$ es el riesgo de muerte de un paciente respecto a otro que tenga una unidad menos de albúmina, en este caso un gramo por litro menos; dicho de otra forma, el paciente con un gramo menos de albúmina en sangre está a $1/0.657=1.52$ veces más riesgo de muerte. El signo negativo del coeficiente de la albúmina nos debe hacer pensar que a mayor cantidad de albúmina el riesgo de muerte descende; la magnitud del descenso viene dada por el valor absoluto del coeficiente. Si el signo del coeficiente hubiese sido positivo indicaría que a mayor nivel del albúmina en sangre, el riesgo de muerte sería mayor.

De forma similar, se puede calcular el riesgo relativo entre dos pacientes cualesquiera; en general este riesgo viene dado por

$$e^{\beta(x_1-x_2)}$$

donde x_1 y x_2 son los valores de la variable en el primer y segundo individuo respectivamente; así, el riesgo de muerte de un individuo con un nivel de albúmina de 34 gr/l respecto de otro con 31 gr/l será $e^{-0.42(34-31)} = e^{-1.26} = 0.28$, es decir, el segundo está a $1/e^{-1.26} = 3.53$ veces más riesgo de muerte que el primero.

Veamos ahora la interpretación de los coeficientes en el caso de que el modelo contenga varias variables predictoras. De forma análoga que en los modelos multivariantes anteriores, en un modelo general con p variables predictoras, el valor e^{β_i} representa el riesgo de muerte entre dos pacientes diferenciados en una unidad de la variable predictora X_i , pero iguales en relación al resto de las variables presentes en el modelo. Al igual que ocurre con otros modelos, la estimación de los parámetros β es complicada y es necesario el uso de programas estadísticos que realicen los cálculos.

La Tabla II.32. representa una estimación del modelo de Cox con factores pronóstico Estadío clínico y Edad, en años, ambas medidas en el momento del diagnóstico, proveniente de un estudio sobre supervivencia de cáncer de mama, en base a una muestra de 300 mujeres procedentes del Registro de Cáncer de Granada.

Tabla II.32.- Estimación del modelo de Cox con factores pronóstico Edad y Estadío.

Variable	Estimación	Error estándar	Riesgo relativo
Edad	0.02	0.009	1.02
Estadío2	0.98	0.442	2.66
Estadío3	1.70	0.433	5.47
Estadío4	3.17	0.501	23.81

¿Cómo interpretar estos resultados? En primer lugar consideremos la edad; en principio, ya que su coeficiente es positivo, indica que a mayor edad de la mujer en el momento del diagnóstico, mayor riesgo de muerte; el valor del coeficiente 0.02 se puede interpretar diciendo que dadas dos mujeres con igual estadío de la enfermedad, si una tiene un año más de edad que otra, la primera está a $e^{0.02} = 1.02$ veces más riesgo de muerte que la segunda; es decir, por cada año de edad, el riesgo de muerte

aumenta en un 2%. Dadas dos mujeres con el mismo estadio, si a la primera se le diagnosticó en cáncer con 45 años y a la segunda con 37, ¿cuál es el riesgo de muerte de la primera respecto de la segunda?; tal riesgo relativo viene dado por $e^{0.02(45-37)} = e^{0.16} = 1.17$, es decir, la mujer diagnosticada a los 45 años está a un 17% de más riesgo que la diagnosticada a los 37, siempre que las dos tuviesen el mismo estadio clínico del tumor cuando éste se les diagnosticó.

La variable estadio clínico necesita un comentario especial; esta es una variable categórica con cuatro niveles: Estadio I, Estadio II, Estadio III y Estadio IV. Como se dijo en II.E.3.1. para el caso que nos ocupa es necesario crear 3 variables indicadoras como sigue: una primera variable, que notaremos por Estadio2, que toma el valor 1 si a la mujer se le diagnosticó un cáncer en estadio II y 0 en caso contrario; una segunda variable, Estadio3, que vale 1 si a la mujer se le diagnosticó el cáncer en estadio III y 0 en caso contrario y, por último, la variable Estadio4 que vale 1 si el estadio del cáncer fue el IV y 0 en caso contrario. Con estas tres variables indicadoras se puede identificar el estadio de cualquier mujer; así, una mujer identificada por la terna (0,0,0) será una a la que se le diagnosticó en un estadio I; (1,0,0) identificará a una mujer con estadio II, (0,1,0) a una en estadio III y, por último, (0,0,1) a una con estadio IV.

Volviendo a la Tabla II.32., ésta es la razón por la que aparecen tres variables que tienen que ver con el estadio; la categoría correspondiente a la terna (0,0,0) es la de referencia y la tabla lo que muestra son los coeficientes de las restantes categorías respecto de ésta; por tanto, el coeficiente 0.98 correspondiente a la variable indicadora Estadio2 tiene la siguiente interpretación: para dos mujeres con igual edad, $e^{0.98} = 2.66$ es el riesgo de muerte de una diagnosticada en estadio II respecto a una diagnosticada en estadio I; $e^{1.70} = 5.47$ es el riesgo de muerte de una diagnosticada en estadio III respecto a una diagnosticada en estadio I y $e^{3.17} = 23.81$ es el riesgo de muerte de una diagnosticada en estadio IV respecto a una diagnosticada en estadio I.

La Tabla II.32., aparte de dar las estimaciones de los parámetros del modelo, de los que acabamos de derivar las estimaciones de los riesgos relativos, también muestra las estimaciones de los errores estándar de tales parámetros, por lo que podremos construir intervalos de confianza. En general, un intervalo de confianza al 95% para un parámetro β_i viene dado por

$$\hat{\beta}_i \pm 1.96 (e.e.)$$

por lo que un intervalo de confianza para el riesgo relativo será

$$e^{\hat{\beta}_i \pm 1.96 (e.e.)}$$

Por ejemplo, un intervalo de confianza para el parámetro correspondiente a la Edad de la mujer será $0.02 \pm 1.96(0.009) = (0.00236, 0.03764)$; por tanto, un intervalo de confianza para el riesgo relativo será $(e^{0.00236}, e^{0.03764}) = (1.002, 1.038)$. Es decir, para mujeres con igual estadio, el riesgo de muerte aumenta como mínimo en un 0.2% y como máximo en un 3.8%, para cada año de diferencia de edad al diagnóstico del tumor; recuérdese que la estimación puntual del aumento del riesgo fue del 2%.

En definitiva, el modelo de Cox estimado para la supervivencia de cáncer de mama que tiene como predictoras a la edad y el estadio al diagnóstico se puede escribir así

$$\hat{h}(t, \text{Edad}, \text{Estadio}) = \hat{h}_0(t) e^{0.02\text{Edad} + 0.98\text{Estadio II} + 1.70\text{Estadio III} + 3.17\text{Estadio IV}}$$

que permite la comparación de dos mujeres cualesquiera. Por ejemplo, supongamos dos mujeres, una con 45 años y un estadio III y otra con 38 años y estadio II; ¿cuál es el riesgo de muerte de la primera respecto de la segunda?. Según la expresión anterior, para la primera mujer

$$\hat{h}(t, 45, \text{Estadio III}) = \hat{h}_0(t) e^{0.02(45) + 0.98(0) + 1.70(1) + 3.17(0)}$$

y, para la segunda

$$\hat{h}(t, 38, \text{Estadio II}) = \hat{h}_0(t) e^{0.02(38) + 0.98(1) + 1.70(0) + 3.17(0)}$$

por lo que el riesgo relativo de la primera respecto a la segunda es el cociente entre estas dos últimas expresiones, es decir,

$$\frac{e^{0.02(45) + 0.98(0) + 1.70(1) + 3.17(0)}}{e^{0.02(38) + 0.98(1) + 1.70(0) + 3.17(0)}} = e^{0.02(45-38) + 0.98(0-1) + 1.70(1-0)} = e^{0.89} = 2.44$$

Decíamos antes que $\hat{h}_0(t)$ es el riesgo estimado de muerte para aquellos pacientes que presenten valores de las variables predictoras igual a 0, en este caso serían individuos con Estadio I (Estadio II=0, Estadio III=0, Estadio IV=0) y Edad=0; la pregunta que surge inmediatamente es: ¿tiene sentido hablar de pacientes con Edad=0? En el caso que nos ocupa no tiene sentido porque el rango de edades de nuestra base de datos no incluye a pacientes recién nacidas diagnosticadas con cáncer de mama y sería absurdo hablar de individuos con 0 años de edad. Para soslayar esta dificultad lo que se suele hacer es la operación que se conoce con el nombre de **centrado** (**cen-**

tering); en nuestro caso, será considerar, en lugar de la Edad, una nueva variable obtenida a partir de ella restándole el valor de alguna medida de posición, por ejemplo la mediana, que en este caso es igual a 60 años, es decir, tomaremos la variable representada por $EdadM=Edad-60$, con lo cual aquellos individuos de 60 años de edad y con Estadío I anulan todas las variables predictoras y formarán el grupo de referencia que da significado al riesgo base. Puede probarse que los coeficientes obtenidos al estimar el modelo con la variable $EdadM$ en lugar de $Edad$ son los mismos y no cambiarán siempre que hagamos transformaciones de las variables iniciales que consistan en sumar o restar constantes, sin embargo los resultados no serán iguales si dividimos o multiplicamos por constantes. En nuestro caso volvemos a tener la misma tabla anterior y el modelo estimado será

$$\hat{h}(t, EdadM, Estadío) = \hat{h}_0(t)e^{0.02(Edad-60) + 0.98Estadío2 + 1.70Estadío3 + 3.17Estadío4}$$

Como cualquier modelo estadístico, el de Cox está basado en una serie de hipótesis que hay que comprobar; este modelo está basado en la denominada hipótesis de **riesgos proporcionales** razón por la que el modelo de Cox también se conoce con el nombre de **modelo de riesgos proporcionales (proportional hazards models)**. Antes se ha visto como el riesgo entre dos sujetos que se diferenciaban en una unidad en términos de la predictora X_i era e^{β_i} , cantidad que es independiente del tiempo, es decir, constante; esta hipótesis es crucial de comprobar siempre que se utilice el modelo de Cox; el ajuste del modelo a nuestros datos, la presencia de observaciones influyentes, la detección de observaciones extremas, etc., son otras cuestiones de interés. De nuevo la colaboración de un estadístico es imprescindible para tratar estas cuestiones; una buena introducción sobre estas materias se puede encontrar en Harris ³⁵.

6. De la asociación a la causalidad

El punto final en cualquier investigación llega cuando tras la recogida de los datos, siguiendo un diseño y análisis adecuados, se obtienen los resultados. La situación puede ser el hallazgo o no de una asociación. Para simplificar imaginemos que el estudio tuvo éxito y se encontró una asociación que es estadísticamente significativa y tiene suficiente significación clínica o práctica. La pregunta que los investigadores quisieran poder contestar ahora es: ¿esa relación es causal? Puede tenerse delante un verdadero factor pronóstico o de riesgo o, lamentablemente, es la consecuencia de una acumulación de errores en el diseño, la recogida de los datos o el análisis.

La respuesta no es en general sencilla. Bradford Hill ³⁶ estableció una serie de condiciones para evaluar evidencias de causalidad que otros autores han incorporado ^{37, 38}

en modelos de etiología multifactorial. En ciencias sociales es difícil encontrar relaciones de causalidad de tipo determinístico. Los problemas de salud están generalmente relacionados con más de una causa, de forma que los distintos determinantes se relacionan entre sí y se producen modificaciones en sus efectos, además de estar relacionados con efectos distintos. Por ello se habla de *urdimbre causal* y de *trama de efectos*³⁹.

De alguna forma ha de hacerse una serie de juicios que permitan tomar la decisión. Elwood⁴⁰ propuso un esquema de evaluación que seguiremos en parte. Se trata de hacer algo semejante a un *diagnóstico de causalidad*: la idea es enfrentar las posibles explicaciones *no causales* a los argumentos que apoyan la posible causalidad en la relación encontrada. No se pretende sólo poner en los platos de una balanza el número de argumentos a favor y en contra, sino ponderar la verosimilitud de cada uno de ellos. Si es posible encontrar más y mejores evidencias de causalidad, entonces se tendrán los elementos suficientes para establecer una conclusión.

6. 1. Validez interna y precisión: errores que pudieran explicar los resultados

Se trata en primer lugar de evaluar la posibilidad de que errores en el diseño (errores sistemáticos o sesgos) o introducidos por el proceso de muestreo (error aleatorio) sean los verdaderos responsables de los hallazgos encontrados en el estudio.

Los **sesgos** posibles pueden afectar, como se ha visto, a la selección de los sujetos estudiados, a la mala clasificación del estatus de exposición o de enfermedad, o a la influencia de factores confundentes. Es difícil encontrar un estudio en el que cualquiera de estos problemas no esté presente. Lo que se pretende es que su influencia sea mínima. Así, es posible aceptar que un estudio esté afectado por una cierta cantidad de sesgo de selección, por ejemplo. Puede ocurrir que un determinado porcentaje de controles se negaran a responder un cuestionario en un diseño de casos y controles. Si se encontró una asociación y los casos presentan más frecuentemente el factor de riesgo, alguien podría argumentar que son precisamente los controles expuestos quienes de forma sistemática se negaron a participar, lo cual provocó el hallazgo de una mayor proporción de casos expuestos.

Una forma de valorar en qué medida ese potencial sesgo de selección ha influido sobre los resultados es ponerse en el caso extremo. Si todos los controles que se negaron a participar hubieran sido expuestos, y se reconstruyera la tabla 2x2 asumiendo éso, ¿cambiaría la estimación de la *odds ratio*? Es decir, ¿desaparecería la asociación, o incluso el factor en estudio podría ser protector? Si puede descartarse esta posibilidad, entonces no hay evidencias de que ese posible sesgo de selección

sea el responsable de la asociación encontrada. De esta forma pueden valorarse cada uno de los sesgos potenciales.

Sin embargo, la influencia de posibles factores confundentes no controlados es un problema sobre todo en estudios no experimentales. Por un lado, el conocimiento profundo del tema en estudio y el apoyo de buenas revisiones bibliográficas permiten decidir qué confundentes controlar, sea en el diseño o en el análisis. Por otro lado, en ausencia de una asignación aleatorizada, siempre pueden existir confundentes desconocidos, e incluso la aleatorización podría no haber conseguido una distribución completamente homogénea de esas variables. Esta amenaza siempre debe tenerse en mente, ya que conclusiones aparentemente claras a día de hoy podrían ser descartadas por futuros estudios.

El **error aleatorio** (grado de precisión) puede controlarse mediante los tests de hipótesis y los intervalos de confianza. En cada situación el investigador debe decidir con qué grado de error se atreve a establecer conclusiones. Si se afirma que existe asociación, el valor de p (error tipo I) informa de la probabilidad de cometer una equivocación. Si el valor de p es muy alto, sería arriesgado hacer tal afirmación. Algo semejante ocurre cuando los intervalos de confianza incluyen o no al valor nulo: si el rango de posibles valores del parámetro no alcanza al valor de *no asociación*, puede establecerse, con una determinada confianza, que la asociación es estadísticamente significativa. Además, para cualquier tamaño de muestra dado es posible determinar la potencia del estudio, es decir, la capacidad de detectar como estadísticamente significativas unas diferencias si es que realmente existen. Sin embargo es más aconsejable que ésto se decida en una fase anterior, calculando previamente el tamaño de muestra mínimo necesario.

Por tanto, el hallazgo de una asociación puede estar producido por algún tipo de error sistemático, por el azar (error tipo I), o porque verdaderamente esa diferencia exista en la población estudiada. El caso contrario, no encontrar asociación estadísticamente significativa, puede estar relacionado con algún sesgo que la esté camuflando, por el azar (falta de potencia de la muestra o error tipo II), o porque realmente esa asociación no existe.

6. 2. Validez interna: argumentos a favor de una explicación de tipo causal

Si el análisis de explicaciones alternativas ha sido satisfactorio, entonces es posible valorar qué evidencias refuerzan la validez interna del estudio en el sentido de apoyar el carácter causal de la asociación encontrada.

Relación temporal.

Es evidente que cualquier relación de tipo causal requiere la presencia de la *causa* (el factor en estudio) antes del *efecto* (la aparición de la enfermedad, muerte, recurrencia,...). En multitud de ocasiones se presentan dificultades, sin embargo. Por ejemplo, el diagnóstico clínico no implica que la fecha de diagnóstico se corresponda con la de aparición de la enfermedad. El periodo de latencia, durante el cual existen alteraciones pero sin expresión clínica clara, puede ser enormemente largo, como en determinados tipos de cáncer. Puede ocurrir que se detecten antecedentes de alguna exposición en una persona y ésta quede clasificada como expuesta, pero la enfermedad se hubiera desencadenado ya cuando ocurrió la exposición. Esta posibilidad hace que en ocasiones se decida no recoger información sobre antecedentes de exposición en los meses o años previos al momento del diagnóstico.

Magnitud de la asociación.

El hallazgo de estimadores de fuerza de asociación (riesgo relativo y otros equivalentes) altos es una protección parcial frente a la influencia de otras explicaciones alternativas. La razón es que asociaciones fuertes difícilmente serían explicables por sesgos, por ejemplo. O, al menos, para que un determinado sesgo pudiera considerarse como la explicación de lo encontrado, tendría que ser un sesgo muy evidente. Por el contrario, una asociación débil podría explicarse por cualquier error o una acumulación sucesiva de errores. Esto explica que muchas asociaciones, si son débiles, sean detectadas por algunos estudios y no por otros.

Relación dosis-respuesta.

La presencia de un gradiente biológico, es decir, a mayor exposición mayor efecto, si no imprescindible, sí es una importante ayuda en la argumentación. La detección de un gradiente se realiza bien comparando diferentes categorías de la exposición con una que se considera basal o de referencia, o bien considerando la variable de exposición como continua, midiendo incrementos en el riesgo por cada aumento de unidad de la variable de exposición. Aunque son posibles relaciones del tipo *todo o nada*, en general es frecuente encontrar relaciones dosis-efecto. Esto puede estar matizado sin embargo por la existencia de niveles umbrales, límites de saturación a partir de los que no aumenta el efecto, y otras situaciones similares.

Cuando en el análisis de los datos se usan modelos matemáticos, un problema frecuente es la aplicación de modelos lineales (regresión lineal, logística,...) sin verificar si la relación entre las variables en estudio es, precisamente, lineal. Puede ocurrir que no sea de esa forma, y el modelo no detecte entonces relación alguna; se puede concluir en ese caso que no existe asociación (*lineal*), de forma errónea.

6. 3. Validez externa: capacidad de generalización de los resultados

La capacidad de extrapolar las evidencias que ofrecen los resultados del estudio a poblaciones distintas a la que aquél analiza depende en gran medida de su precisión y validez interna. Sin embargo es frecuente enfrentarse al dilema de si aumentar la validez interna a costa de disminuir la validez externa. Por ejemplo, la existencia de un factor confundente previamente conocido (edad, sexo, otros factores de riesgo, etc) puede motivar la aplicación de algún criterio de exclusión (por ejemplo, no incluir a mujeres, o a hombres). La neutralización de un factor confundente puede traer como consecuencia un aumento en la validez interna. Sin embargo, surge entonces la pregunta: ¿son los resultados y conclusiones aplicables a la población excluida del estudio?

Por otro lado, la extrapolación de los resultados a otras poblaciones distintas de la población diana en la que se aplicó el estudio suele realizarse mediante procedimientos de inferencia estadística, no siempre debidamente justificados. Aún cuando pudiera haberse analizado a toda la población diana en lugar de muestrearla, los investigadores se enfrentan a un nuevo problema: ¿es esa población estudiada representativa de alguna otra que no es posible definir?. Existe controversia en este punto, pero una solución *general* podría apuntar en el sentido de dar más valor a los requisitos de causalidad que el estudio aporta, y mantener un nivel de prudencia al interpretar los resultados de las pruebas de inferencia estadística cuando es difícil definir la *población* a la que se pretende inferir las conclusiones.

6. 4. Comparación de los resultados del estudio con evidencias externas

Un punto final importante en cualquier discusión es la relativización de los resultados en función de lo conocido a partir de otras fuentes.

Especificidad.

La relación exclusiva entre un factor y un efecto, y no con otros, es uno de los puntos propuestos por Hill que más críticas ha despertado. Conforme avanza el conocimiento sobre los problemas de salud y sus determinantes, cada vez es más difícil encontrar situaciones de este tipo; más frecuentemente la situación es de multicausalidad, y además una misma exposición suele estar implicada con enfermedades diferentes.

Plausibilidad biológica.

Una buena capacidad de comprensión de los mecanismos subyacentes a la asociación encontrada refuerza la argumentación sobre su carácter causal. Evidentemente ésto es más difícil conforme más se aleje la hipótesis en evaluación del mundo

físico, bioquímico, etc. Sin embargo, hallazgos procedentes de estudios en un nivel más social o sociológico suelen marcar pistas para la investigación ulterior. La capacidad del conocimiento actual de ofrecer explicaciones plausibles puede ser insuficiente, y ésto es una limitación. Las hogueras de la Inquisición pudieron ser la consecuencia de esta limitación en más de una ocasión.

Consistencia.

Resultados similares encontrados en estudios llevados a cabo en otras poblaciones, frecuentemente con distintos diseños, son uno de los argumentos más frecuentemente utilizados en la discusión de cualquier publicación científica. La principal razón, aparte del razonable alivio, es que una asociación que detectan diseños con sesgos probablemente distintos, en poblaciones en las que la distribución de distintos confundentes puede ser variada, y en circunstancias diferentes, es una asociación más verosímil. Se ha sugerido que la consistencia podría también verificarse en distintas subpoblaciones dentro del propio estudio; sin embargo esto es frecuentemente difícil, en especial porque el tamaño de la muestra no suele calcularse con ese propósito.

Coherencia.

La coincidencia de los resultados del estudio con la distribución del factor de exposición en ésa y otras poblaciones, y la correspondencia con la distribución del efecto, es otro argumento a confrontar. Si la asociación es de tipo causal, se espera entonces que en poblaciones con mayores niveles de exposición la enfermedad sea más frecuente. Paradójicamente el último punto presentado para la valoración de una relación causal acude en ayuda de las evidencias que se presentaron como más débiles cuando se revisaron los distintos tipos de diseños, las procedentes de estudios ecológicos.

F. ANÁLISIS DE LA DECISIÓN CLÍNICA

J.A. Chicharro Molero

R. Burgos Rodríguez

1. Introducción. Variabilidad

Posiblemente si todas las decisiones que los clínicos tomamos, o al menos la mayoría, no tuvieran ningún grado de incertidumbre, nunca se hubiera desarrollado toda una metodología para ayudar al clínico en aquellas decisiones de casos difíciles, los típicos casos de discusión en sesión clínica, o para la construcción de algoritmos de diagnóstico-tratamiento que definan una línea de actuación común.

Desde luego, sólo es necesario un mínimo ejercicio de reflexión para darse cuenta de que la decisión clínica está sujeta a variabilidad por varias circunstancias:

- 1) En algunos casos es necesario decidir sobre la aplicación de algún tratamiento o la realización de algún test diagnóstico sin que se conozca con exactitud, por que el conocimiento no ha avanzado; por ejemplo un varón de 64 años con síntomas miccionales, tacto rectal normal, PSA de 7 ng/ml y ecografía sin hallazgos, **¿qué hacer?**; seguramente algunos médicos urólogos realizarían biopsia transrectal⁴¹ y otros evaluarían el PSA densidad, incluso podría haber otra tercera postura que sería evaluar al año el PSA velocidad, pero es que podríamos optar por una cuarta, evaluar el PSA por rangos de edad⁴². **¿Cuál sería la actitud más correcta y, sobre todo, mejor para el paciente?**

En este primer punto observamos cómo la variabilidad se debe a que la práctica médica en ocasiones nos obliga a tomar decisiones en situaciones que no están totalmente consensuadas.

- 2) La decisión puede cambiar, así mismo, porque la información recibida varíe de unos casos a otros debido a la realización de exploraciones complementarias que son operador dependiente, como por ejemplo la ecografía.

- 3) También la decisión puede cambiar de unas zonas geográficas a otras; en un estudio realizado por Medicare en 16 comunidades que pertenecían a cuatro estados en E.E.U.U., se pudo apreciar que en Iowa el 15% de los varones menores de 85 años habían sufrido cirugía prostática mientras que en otro estado el 60% de los varones por debajo de esa edad estaban operados; también la tasa de histerectomías realizadas en mujeres menores de 70 años en Maine variaba del 20% al 70% dependiendo de la comunidad; en Vermont la probabilidad de encontrar un niño amigdalectomizado oscila del 8% al 70% dependiendo de la comunidad. Sin duda estas variaciones podrían explicarse, en parte, por diferentes incidencias de enfermedades, recursos disponibles o, incluso, preferencias del paciente, pero parece poco creíble que estas circunstancias expliquen toda esta variabilidad.

En definitiva existe variabilidad en la toma de decisiones que cotidianamente realizamos los médicos clínicos; la asunción de esta aseveración tiene implicaciones de diversa índole ya que esto querrá decir que las decisiones que individualmente hacemos los médicos no son siempre necesariamente correctas. Por lo tanto los pacientes deberían pedir una segunda opinión y también se verían afectados conceptos como el consentimiento informado, opinión de expertos, desarrollo de consensos, definición de estándares, malpraxis y calidad asistencial.

Por lo tanto, aunque esta incertidumbre en la decisión clínica no se produce en todas las ocasiones ni cotidianamente, debería poder cuantificarse de tal manera que el clínico posea una herramienta metodológica en aquellas situaciones que las requiera. Para soslayar esta cuestión, desde la última década se han incorporado a la medicina unas técnicas agrupadas bajo el epígrafe de «**análisis de la decisión**»^{43, 44} que ya se venían utilizando desde hace bastante más tiempo en otras áreas como el marketing, economía o en el desarrollo de ingenios militares.

2. Arbol de decisión

Es un método que permite representar en forma de árbol todas las alternativas de cualquier problema que surja, cuantificando y valorando cual de ellas es mejor dependiendo de valoraciones de resultados diferentes como supervivencia en años sola o ajustada por calidad de vida, utilidad, o introduciendo términos económicos y pasando a coste-efectividad, coste-utilidad o, de una forma aún más complicada, como es el coste-beneficio.

Para continuar con el desarrollo de esta metodología nos basaremos en un ejemplo práctico que iremos desarrollando, motivo de alguna controversia actual en el campo de la Urología:

Varón de 65 años que acude a consulta por prostatismo de 24 meses de evolución; IPSS predominantemente obstructivo con un puntaje de 18 con calidad de vida 3 (L-3). Tacto rectal normal. PSA de 21 ng/ml. La ecografía transrectal muestra un nódulo hipoecoico de 2,5 cc en lóbulo derecho. Se practica biopsia transrectal ecodirigida, hallándose Adenocarcinoma de próstata en ambos lóbulos con valor patrón de Gleason 6. Ante este hallazgo se decide Cirugía radical por vía retropúbica; la cuestión es decidir si la linfadenectomía se debe hacer por vía laparoscópica o laparotómica.

Para construir un árbol de decisión sobre este problema o cualquier otro más complejo hay que seguir una serie de pasos ^{43, 44}:

- 1) Construcción del árbol de decisión.
- 2) Asignación de probabilidades a cada rama del árbol que lo requiera.
- 3) Valoración de los resultados finales en los términos que decidamos.
- 4) Calcular cual es el valor esperado de las alternativas del nodo raíz, estableciendo cual de ellas es la dominante, la de mayor valor esperado.
- 5) Observar si la alternativa dominante sigue siendo la misma cuando algunas probabilidades y/o resultados de las ramas del árbol oscilan en un rango de valores clínicamente plausible. Esto se conoce como **análisis de sensibilidad**.

2.1. Construcción del árbol de decisión

Para construir un árbol hay que cumplir dos requisitos: conocer el método y ser experto en la disciplina que se aplica éste. Las ramas del árbol de decisión parten siempre de lo que se denomina **nodos del árbol** (figura II.30) que pueden ser de tres tipos: nodos de decisión, nodos probabilísticos y nodos terminales.

Los **nodos de decisión** se representan por un cuadrado (figura II.30) y todo árbol debe empezar por él, sin perjuicio de que a lo largo del desarrollo pueden aparecer más o no. Este nodo supone un punto donde el clínico puede decidir que hacer, de ahí su nombre; en nuestro ejemplo hemos decidido hacer linfadenectomía laparoscópica o abierta (figura II.30), pero podríamos haber incluido otra tercera opción, por

ejemplo no hacer linfadenectomía, pero en nuestro caso el problema que se nos plantea es cuál de los dos tipos de linfadenectomías es más útil.

Los **nodos probabilísticos o aleatorios** se representan por un círculo (figura II.30) y son puntos en el árbol donde se producen hechos que el clínico no puede decidir que ocurre, sino que suceden con una determinada probabilidad; en nuestro ejemplo de la figura II.30 el urólogo decide hacer linfadenectomía laparoscópica o no, pero una vez que hace una u otra puede ocurrir que los ganglios sean positivos o negativos con una determinada probabilidad y esto escapa a la voluntad del clínico. De cada nodo probabilístico deben salir tantas ramas como sucesos posibles pueden darse, en nuestro caso sólo dos, que deben ser incompatibles.

Los **nodos terminales** representan el final de cada rama del árbol y tendrán unos valores dependiendo de cómo queramos evaluar el árbol; en nuestro caso lo haremos en términos de utilidad, aunque también se podría hacer en términos de coste-utilidad.

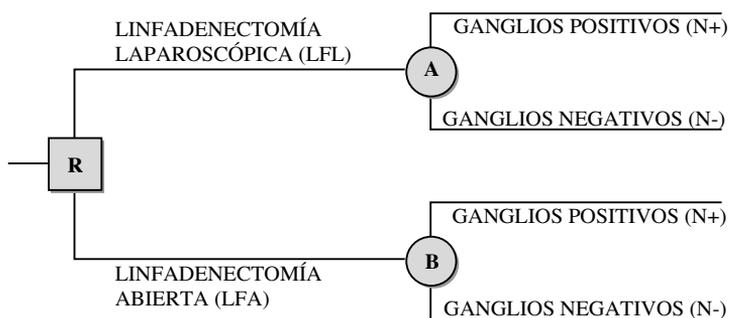


Fig. II.30.- Representación gráfica de un árbol de decisión en su primera fase.

Una vez que hemos definido el problema, las alternativas iniciales y los elementos del árbol continuaremos con la construcción del mismo; para esto es interesante describir todas las posibles alternativas después de cada rama y si éstas ocurrirán aleatoriamente, y por tanto saldrán de un nodo probabilístico, o por el contrario serán decisión del clínico y saldrán de un nodo de decisión. En nuestro ejemplo después de los resultados puede ocurrir que haya complicaciones de la técnica o no las haya y esta circunstancia no depende de la voluntad del clínico por lo que habrá otro nodo probabilístico al final de cada una de las ramas del nodo A y B de la figura II.30.; de cada uno de los nuevos nodos probabilísticos que denominaremos A₁, A₂, B₁ y B₂ saldrán dos nuevas ramas, figura II.31., que serán la presencia o no de complicacio-

nes tras linfadenectomía laparoscópica, nodos A₁ y A₂, y tras linfadenectomía abierta, nodos B₁ y B₂. El árbol en esta fase quedará como se puede apreciar en la figura II.31.

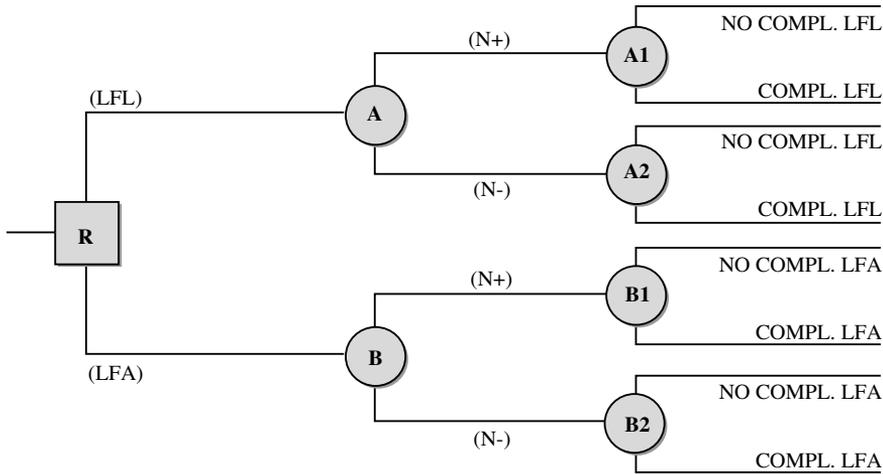


Fig. II. 31.- Representación gráfica de un árbol de decisión en su segunda fase.

Otro de los requisitos importantes a la hora de construir un árbol de decisión es fijarse cual debe ser el horizonte temporal del problema que nos planteamos, en nuestro caso se trataba de ver que técnica de linfadenectomía era más útil en un paciente que sería sometido a cirugía radical. Continuando con la construcción del árbol y observando la figura II.31 apreciamos como de las ocho ramas, 4 de ellas, las originadas en los nodos A₁ y B₁ tienen afectación ganglionar, por lo que estas ramas terminarán aquí, otro autor podría operar en estos casos o introducir un nodo de decisión y operar o no dependiendo de la edad y número de ganglios afectados, pero esto sería para un árbol de propósito más general, por lo que en nuestro supuesto continuamos con la tesis de que cuando hay afectación ganglionar no se hace cirugía radical, por lo que las ramas de los nodos A₁ y B₁ han llegado a su fin. Sin embargo las ramas de los nodos A₂ y B₂ son supuestos sin afectación ganglionar por lo que en este caso se realizará una prostatectomía radical retropúbica y al realizar esta técnica quirúrgica podrán aparecer complicaciones o no independientes de la voluntad del cirujano por lo que habrá nodos probabilísticos que denominaremos A₂₁, A₂₂, B₂₁ y B₂₂, figura II.32., saliendo de cada uno de ellos dos ramas que contemplan la posibilidad o no de complicaciones tras prostatectomía radical retropúbica, quedando el árbol como se aprecia en la figura II.32. y que sería el árbol completo de nuestro supuesto.

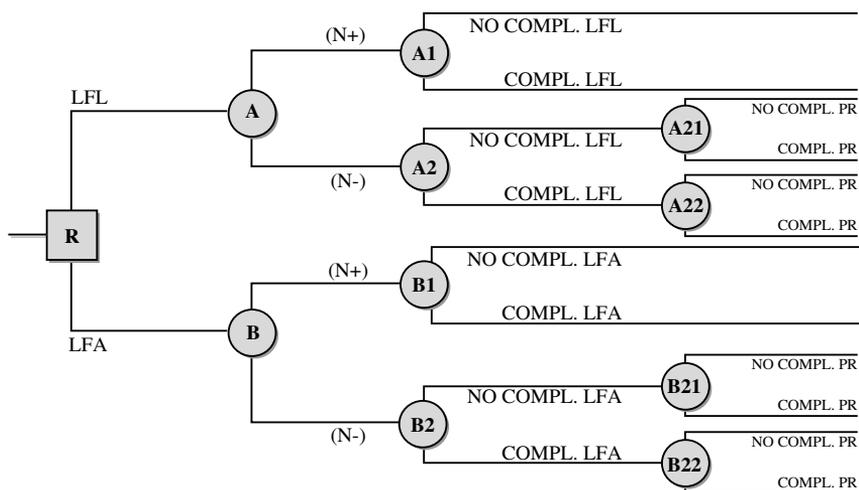


Fig. II. 32.- Representación gráfica de un árbol de decisión en su última fase.

En resumen para la construcción del árbol hay que identificar el problema, limitar el horizonte temporal e ir desarrollando de izquierda a derecha la secuencia de sucesos lógicos que pueden ocurrir, determinando si partirán de un nodo probabilístico o de decisión.

2.2. Asignación de probabilidades

Este paso en el análisis de decisión es uno de los fundamentales pues de los valores numéricos que asignemos dependerá cual de las ramas del nodo raíz (primer nodo de decisión) será la alternativa dominante. Sólo a las ramas que salen de nodos probabilísticos se les asignará probabilidades. Esta asignación es una de las tareas más complejas y prolijas en todo el análisis de la decisión pues las respectivas probabilidades, por ejemplo de complicaciones tras linfadenectomía laparoscópica, se pueden extraer de dos fuentes: la experiencia propia o la literatura, pero en el primer supuesto debemos estar seguros de que nuestra experiencia no esté sesgada o sea pequeña, y en el segundo debemos testar las fuentes de la literatura y contrastar si los datos han sido obtenidos con buena metodología (¡¡¡¡¡hay que aprenderse esta monografía!!!!); pues, aunque podríamos asignar números de cualquier fuente, no debemos olvidar que la solidez de nuestro análisis de decisión va a depender, entre otros factores, de la asignación de probabilidades. También hemos de decir que es difícil en medicina que cualquier probabilidad tenga un valor exacto y siempre habrá un rango de valores, pero este tema lo abordaremos más adelante en el análisis de sensibilidad.

En nuestro ejemplo, al no ser extremadamente complicado, la asignación de probabilidades, una vez conocidas, no entraña mucha complejidad, pero en otros casos

sobre todo cuando hay test diagnósticos, en el árbol habrá probabilidades que sean los valores predictivos o probabilidades que aunque no estén en la literatura se puedan calcular a partir de otras.

Consideraremos que el porcentaje de complicaciones será 8% , 4% y 16% para la linfadenectomía laparoscópica, linfadenectomía abierta y prostatectomía radical ⁴⁵; la probabilidad de afectación ganglionar en nuestro caso será del 42%. Estos valores los trasladaremos a sus respectivas ramas del árbol en términos de probabilidad, sólo hay que dividirlo por 100, resultando el árbol de la figura II.33..

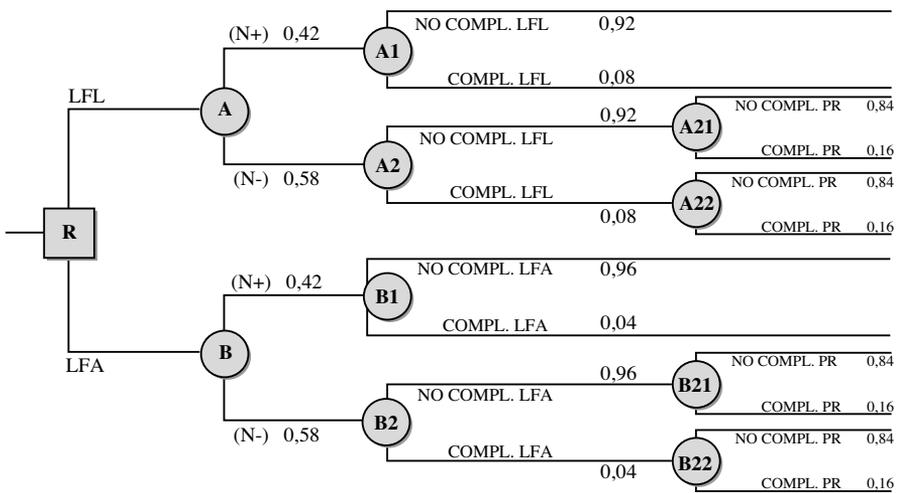


Fig. II. 33.- Asignación de probabilidades a las ramas de un árbol de decisión.

2.3. Valoración de los resultados

Otra de las cuestiones importantes es en qué términos valoraremos el árbol, pues de ello puede depender también la alternativa dominante del mismo. Los más utilizados son efectividad y utilidad, aunque también se usan otras valoraciones como coste-efectividad, coste-utilidad y coste-beneficio.

La **efectividad** es la valoración que más acostumbrados estamos a hacer los clínicos, pero no siempre es la más adecuada; usualmente se puede valorar en supervivencia o en porcentaje de curación. Pero muchas veces hay que plantearse si el aumentar la supervivencia un poco más de tiempo merece la pena si a cambio empeora la calidad de vida, sobre todo habría que preguntarse o preguntarle al paciente si esto merece la pena; por eso en los últimos años ha surgido el término de «*supervivencia en años ajustada por calidad de vida*». El problema surge en cómo evaluar la calidad de vida que no trataremos aquí por ser un tema que se escapa al contenido de este libro,

pero hemos de advertir que los «qualy» están basados en cuestionarios específicos para cada situación y es resbaladizo y poco adecuado utilizar cuestionarios en circunstancias para las que no están diseñados ni validados.

La **utilidad** es una valoración poco familiar en el ámbito clínico; permite tener en cuenta la opinión del paciente en la valoración del resultado, y por tanto, en cierto sentido, es una medida de la preferencia del paciente por el resultado. Para el cálculo de la utilidad se han de seguir varios pasos:

1. Listar todos los posibles resultados que se puedan dar en nuestro árbol ya construido, y disponerlos por orden de preferencia.
2. Al mejor resultado se le dará el valor 1 y al peor el valor cero, ya que la utilidad oscila entre estos dos valores.
3. Al resto de posibles resultados intermedios se le asignan valores entre 0 y 1 dependiendo de las preferencias del paciente.

Pero, como resulta lógico, establecer la utilidad en una entrevista para resolver su caso es bastante difícil, por no decir extremadamente difícil, por lo que normalmente se utilizan otras valoraciones de resultados.

Los términos que incluyen el coste obligará a calcular el precio de cada uno de los procedimientos o test diagnósticos que se encuentren en el árbol, en nuestro ejemplo deberíamos saber cuanto cuesta una linfadenectomía laparoscópica, una linfadenectomía abierta y una prostatectomía radical. Cada nodo terminal tendrá un precio que equivaldrá a la suma de los precios de los procedimientos que se hayan realizado en el trayecto que une ese nodo terminal con el nodo raíz. Cuando se valora el árbol en dos términos como coste-utilidad o coste-efectividad se deberá calcular el valor esperado de cada alternativa del nodo raíz para cada una de las valoraciones, es decir en dinero y en efectividad o utilidad, pero esto lo desarrollaremos en el siguiente apartado.

Nuestro árbol será valorado en términos de utilidad; el cálculo de los valores de la misma puede encontrarse en el artículo de Wolf⁴⁵.

2.4. Cálculo del valor esperado

Una vez llegados a este punto, sólo nos queda saber sobre qué alternativa del nodo raíz, linfadenectomía laparoscópica (LPL) o linfadenectomía abierta (LFA), nos inclina el árbol de decisión construido.

Si el árbol se construya de izquierda a derecha, se evalúa al revés, de derecha a izquierda. La metodología consiste en ir calculando los valores esperados en cada nodo del árbol comenzando por la derecha hasta llegar a los iniciales, obteniendo los valores esperados de las alternativas de decisión, LFL y LFA, que en nuestro caso son en términos de utilidad. Para facilitar la comprensión del texto nos referiremos al árbol de la figura II.34 que es el mismo de la Fig. II.33., pero con las utilidades de cada nodo terminal y quitando la denominación alfabética de las ramas del árbol.

El valor, utilidad, esperado en un nodo probabilístico es igual a la suma de los productos obtenidos de multiplicar la probabilidad de sus ramas por la valoración del resultado en cada una de ellas; así la utilidad esperada para el nodo A₁ será:

$$1 \times 0.92 + 0.9 \times 0.88 = 0.99$$

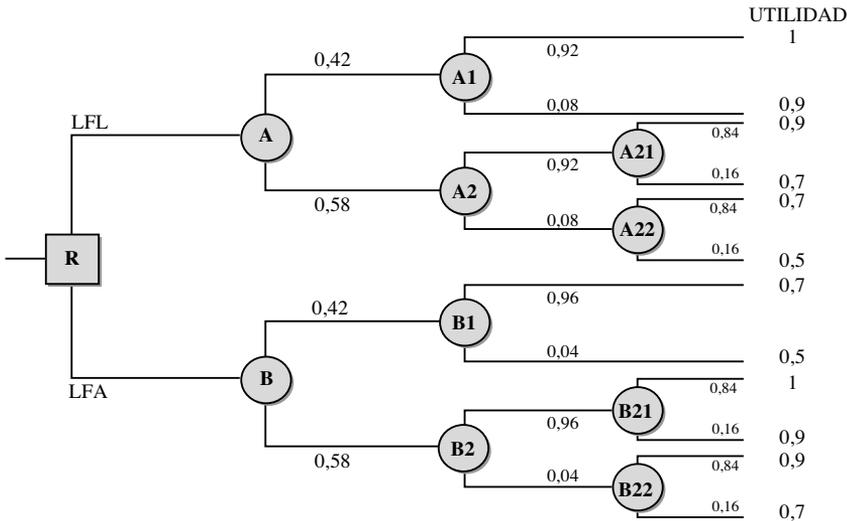


Fig. II. 34. Asignación de la utilidad a las ramas de un árbol de decisión.

De igual forma las utilidades esperadas para los nodos A₂₁ y A₂₂ serán respectivamente:

$$0.9 \times 0.84 + 0.7 \times 0.16 = 0.86$$

$$0.7 \times 0.84 + 0.5 \times 0.16 = 0.57$$

Siguiendo el mismo razonamiento las utilidades esperadas de los nodos B₁, B₂₁ y B₂₂ son 0.69, 0.98 y 0.86, respectivamente. La cuestión que surge es cómo se calcula las utilidades esperadas de los nodos A₂ y B₂; pues bien las probabilidades de las ramas de ambos nodos está claro cuales son pues están representadas en la figura II.34. y las utilidades esperadas que se deben atribuir a cada rama son las calculadas para los nodos a los que llegan éstas, así la de la rama superior e inferior del nodo A₂ serán las

calculadas para los nodo A_{21} y A_{22} ; y de igual manera las utilidades de la rama superior e inferior del nodo B_2 , serán las calculadas para los nodos B_{21} y B_{22} , respectivamente.

Los cálculos para lo nodos A_2 y B_2 quedarían como sigue:

$$0.86 \times 0.92 + 0.57 \times 0.08 = 0.83 \qquad 0.98 \times 0.96 + 0.86 \times 0.04 = 0.97$$

De igual forma se calcularían las utilidades esperadas para los nodos A y B

$$0.99 \times 0.42 + 0.83 \times 0.58 = 0.89 \qquad 0.69 \times 0.42 + 0.97 \times 0.58 = 0.85$$

que corresponden, respectivamente a las utilidades de la LFL , 0.89, y LFA, 0.85; por tanto y atendiendo a nuestro análisis de decisión y los valores obtenidos en el árbol de nuestro modelo, en el paciente del caso presentado con una probabilidad de afectación ganglionar del 42% deberíamos practicar una linfadenectomía laparoscópica y posterior prostatectomía radical, si procede.

2.5. Análisis de Sensibilidad

Pero si el análisis de la decisión clínica terminara en el apartado anterior, no sólo resultaría pobre, sino también polémico, pues seguro que hay autores que sostienen, mediante bibliografía o análisis de la experiencia propia, que las complicaciones de la cirugía laparoscópica son mayores del 8%,o menores. De igual manera ocurrirá con los porcentajes de complicaciones que hemos manejado para la linfadenectomía abierta y para la prostatectomía radical.

Para evitar tal controversia existe el *análisis de sensibilidad* que consiste en evaluar un árbol por cada uno de los valores del rango en el que pensemos que se mueve una probabilidad. Es decir, si creemos que las complicaciones de la prostatectomía radical oscilan de un 8% a un 32% deberemos construir tantos árboles como valores intermedios queramos evaluar de esta probabilidad. Asimismo, si pensamos que otras probabilidades no son un número fijo, sino que oscilan dentro de un rango deberemos evaluar un árbol por cada una de las combinaciones de las diferentes probabilidades, pero ¡¡¡no hay que desesperarse!!!!, pues «*la ciencia adelanta que es una barbaridad*» y la informática ha resuelto este problema ya que existe software que, una vez contruido el árbol, asignadas las probabilidades y la valoración de resultados, con una leve pulsación de la tecla [RETURN] evalúa en décimas de segundo un árbol, permitiendo introducir en las ramas no sólo probabilidades, sino también rangos de ellas.

Asimismo podemos efectuar nuestro análisis de decisión, haciendo que los resultados no sean cifras fijas, sino rangos razonables, que sean plausibles en clínica.

Si después de todas estas variaciones, la alternativa de decisión con mayor valor, utilidad, esperado continua siendo la misma, diremos que nuestra decisión es insensible a los cambios.

3. Probabilidad umbral

Pero la decisión analizada por el árbol de nuestro ejemplo era para cualquier paciente que tenga una probabilidad de afectación ganglionar del 42%, pero ¿qué ocurre si esta probabilidad aumenta o disminuye?, y lo que es más importante ¿a partir de qué probabilidad de afectación ganglionar la utilidad de la linfadenectomía laparoscópica es mayor que la utilidad de la linfadenectomía abierta?

Para contestar a esta interesante pregunta deberíamos calcular las utilidades de LFL y LFA para cada uno de los valores de la probabilidad de afectación ganglionar; representaremos en un eje cartesiano (X,Y) los valores, en el eje Y la utilidad esperada y en el X la probabilidad de afectación ganglionar o el porcentaje, los valores de nuestro ejemplo se aprecian en la figura II.35., observando que hay un punto donde la utilidad de LFL y LFA son idéntica, denominándose a la probabilidad de afectación ganglionar que corresponde a este punto, probabilidad umbral, 0.39 (39%) en nuestro ejemplo. Quiere esto decir que en cualquier caso, con una probabilidad de afectación ganglionar superior al 0.39 (39%), la utilidad de la linfadenectomía laparoscópica es superior a la de la linfadenectomía abierta.

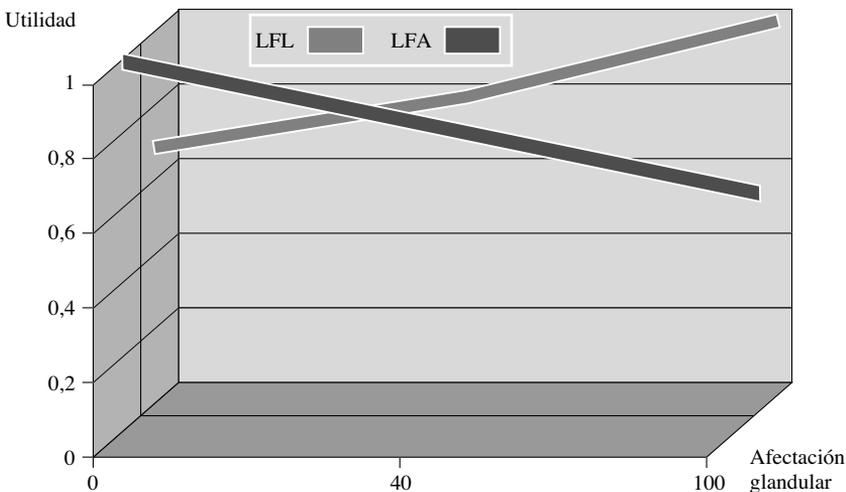


Fig. II.35.- Representación gráfica de la afectación ganglionar y utilidad.

REFERENCIAS

1. Moses LE. Statistical concepts fundamental to investigations. En: Bailar III JC, Mosteller F eds. *Medical Uses of Statistics*, Second Edition. Boston: The New England Journal of Medicine Books, 1992.
2. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. En: Bailar III JC, Mosteller F eds. *Medical Uses of Statistics*, Second Edition. Boston: New England Journal of Medicine Books, 1992.
3. Instructions to authors. *British Medical Journal* 1994; 308; 39-42.
4. Andersen B. *Methodological errors in medical research. An incomplete catalogue*. Oxford: Blackwell, 1990.
5. Altman D. The scandal of poor medical research. *British Medical Journal* 1994; 308: 283-284.
6. Tufté ER. *The visual display of quantitative information*. Cheshire, Conn.: Graphics Press, 1983.
7. Sanchez-Cantalejo Ramírez E. *Test Diagnósticos*. Documento de Apoyo del Curso Intermedio de Epidemiología y Estadística. Granada: Escuela Andaluza de Salud Pública, 1991
8. Timmis AD. Probability analysis in the diagnosis of coronary disease. *Br Med J*. 1985; 286:281
9. Sackett DL, Haynes RB, Guyatt GH y cols. *The interpretation of Diagnostic Data*. En: Sackett DL, Haynes RB, Guyatt GH y cols. *Clinical Epidemiology: A basic science for clinical medicine*. Second edition. Boston: Little Brown and company, 1991:69-151
10. Spicer CC. Test Reduction: II-Bayer's theorem and the evaluation of tests. *Br Med J*. 1980; 281-592
11. Metz CE. *Basic principles of ROC analysis*. Semin Nuclear Med. 8-283. 1978
12. Thompson ML, Zuchini W. On the statistical analysis of ROC curves. *Statist Med*. 1988; 8:1277
13. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology*. 1982; 143:29
14. Gardner MJ, Altman, DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal* 1986; 292: 746-50.
15. Mora MJ, Extramiana J, Navarro FJ et al. Flurbiprofen: Alternativa terapéutica en el dolor del cólico nefrítico. *Arch. Esp. de Urol* 1992; 45(8): 773-7.

16. Ware J H, Mosteller F, Delgado F et al. P values. En: Bailar III JC, Mosteller F eds. *Medical Uses of Statistics*, Second Edition. Boston: The New England Journal of Medicine Books, 1992: 181-200.
17. Matthews DE, Farewell VT. *Using and understanding Medical Statistics*, Second Ed.. Basel: Karger, 1988.
18. Freiman JA, Chalmers TC, Smith H Jr. et al. The importance of beta, the type II error, and sample size in the design and interpretation of the randomized controlled trial: survey of 71 «negative» trials. *N Engl J Med* 1978; 299: 690-4.
19. Schlesselman JJ. *Case-Control Studies*. New York: Oxford University Press, 1982.
20. Machin D, Campbell MJ. *Statistical Tables for the Design of Clinical Trials*. Oxford: Blackwell Scientific Publications, 1987.
21. Instructions to authors. *British Medical Journal*. 1994; 308: 39-42.
22. Altman DG. Statistics and ethics in medical research. Misuse of statistics is unethical. *British Medical Journal*. 1980; 281: 1182-4.
23. Armitage P, Berry G. *Statistical Methods in Medical Research*, 3th ed. Oxford: Blackwell, 1994.
24. Kramer MS. *Clinical Epidemiology and Biostatistics*. New York: Springer Verlag, 1988.
25. Altman DG. *Practical statistics for medical research*. Chapman and Hall: London, 1991.
26. Miller RG. *Beyond Anova, basics of applied statistics*. New York: John Wiley, 1986.
27. Brand R. Using logistic regression in perinatal epidemiology: an introduction for clinical researchers. Part 2: the logistic regression equation. *Paediatric and Perinatal Epidemiology* 1990;4:221-235.
28. Catalona WJ, Smith DS, Ratliff TL et al. Measurement of prostatic-specific antigen in serum as a screening test for prostate cancer. *New Engl J Med* 1991; 324:1156-1161.
29. Machin D, Campbell MJ. *Statistical tables for the design of clinical trials*. Oxford: Blackwell, 1987.
30. Fleiss, JL. *Statistical methods for rates and proportions*. 2nd edn. New York: Wiley, 1981.
31. Pringle M, Stewart-Evans C, Coupland C et al. Influences on control in diabetes mellitus: patient, doctor, practice, or delivery of care?. *British Medical Journal* 1993; 306:630-634.
32. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health* 1989;79:340-349.
33. Horbar JD, Onstad L, Wright E, et al. Predicting mortality risk for infants weighing 501 to 1500 grams at birth: A National Institutes of Health Neonatal Research Network report. *Critical Care Medicine* 1993; 21: 12-18.
34. Christensen E. Multivariate survival analysis using Cox's regression model. *Hepatology* 1987; 7:1346-1358.

35. Harris EK, Albert A. *Survivorship analysis for clinical studies*. New York: Marcel Dekker Inc., 1991
36. Hill AB. The environment and disease; association or causation. *Proc R Soc Med* 1965;58:295-300.
37. Rothman K. Causes. *Am J Epidemiol* 1976;104:587-592.
38. Miettinen O. *Theoretical Epidemiology: principles of occurrence research in Medicine*. New York: Wiley, 1985.
39. MacMahon B, Pugh TF. *Epidemiology, Principles and Methods*. Boston Mass.: Little, Brown & Co., 1970.
40. Elwood JM. *Causal relationships in Medicine. A practical system for critical appraisal*. Oxford: Oxford Univ Press, 1988.
41. Catalona WJ, Smith DS, Ratliff TL et al. Measurement of prostate specific antigen in serum as a screening test for prostate cancer. *N Eng J of M*. 1991; 324 (27):1156-1161
42. Oesterling JI, Jacobsen T, Chute CG et al. Serum prostate specific antigen in a community based population of healthy men: establishment of age specific reference ranges. *JAMA*. 1992; 270:860-864.
43. Sánchez JJ. *Análisis de decisión. Documento de apoyo del curso Análisis de decisión en Clínica*. Granada: Escuela Andaluza de Salud Pública, 1993.
44. Sackett DL, Haynes, RB, Guyatt GH y cols. *The interpretation of Diagnostic Data*. En: Sackett DL, Haynes, RB, Guyatt GH et al. *Clinical Epidemiology: A basic science for clinical medicine*. Second edition. Boston: Little, Brown and company, 1991:69-151.
45. Wolf JS, Perinchery N, Katsuto M et al. Laparoscopic pelvic lymphadenectomy criteria for patient selection. *Urology*. 1993; 42:6.

PARTE III
EL SOPORTE INFORMÁTICO EN LA INVESTIGACIÓN

A. INTRODUCCIÓN

A. Calzas Urrutia

Se suele decir frecuentemente que la investigación, en cualquiera de sus fases, requiere no sólo de rigor metodológico sino además de un cierto dominio de toda una serie de técnicas instrumentales que van a ser utilizadas en alguna de las fases del proyecto de investigación. En concreto, en la Investigación Sanitaria (en general, en las Ciencias Sociales) cuando se opta por "la vía cuantitativa" se hace necesaria con frecuencia la utilización de las técnicas informáticas de tratamiento de datos.

Es aquí donde empiezan los problemas para muchos. Las actitudes hacia la utilización personal de la informática se muestran en un amplio abanico: pavor hacia las "maquinitas", recelo, bloqueo mental, complejo de inferioridad por no saber utilizar algo que hoy en día "lo saben hasta los niños", ... son entre otras muchas las actitudes que hemos captado con frecuencia.

Por si fuera poco, se dice además, que hoy día hay que ser trilingüe: además del idioma materno y el inglés hay que saber informática. Y es que el lenguaje que subyace en la Informática ha evolucionado tan rápidamente como ésta misma. Parece como si ya una barrera inicial se presentara al usuario que necesita de los servicios de la informática para su proyecto personal. El lenguaje, la comunicación al fin y al cabo, está de fondo.

¿Es la Informática un campo sólo para los profesionales de la misma? Afortunadamente, no. Los informáticos no han olvidado la figura (a veces "desvalida") del usuario¹ y han desarrollado una serie de programas de tratamiento de datos no demasiado complejos de utilizar.

Es nuestro propósito en estas páginas dar una leve idea de cómo el usuario "profano" puede hacer un buen uso de la Informática. Para ello, daremos un repaso al proceso de tratamiento de datos, desde que se diseña un protocolo de recogida de datos hasta que se produce la obtención de información, haciendo hincapié en toda una serie de factores que hay que tener en cuenta y dando referencias de ciertos programas que nos pueden ayudar a resolver nuestra tarea.

¹ El usuario se define como "persona a quien se considera que la informática rinde sus servicios". Castelani, X. Método General de Análisis de una Aplicación Informática. Barcelona: Ed Masson, 1986

B. INFORMACION Y DATOS: PREPARACION DE LOS DATOS

A. Calzas Urrutia

1. Introducción

Se dice que un ordenador es una máquina capaz de aceptar una entrada automatizada de datos, procesar de forma mecánica y lógica estos datos y ofrecer una salida, asimismo automática, del resultado de este procesamiento¹.

La figura siguiente esquematiza el proceso²:

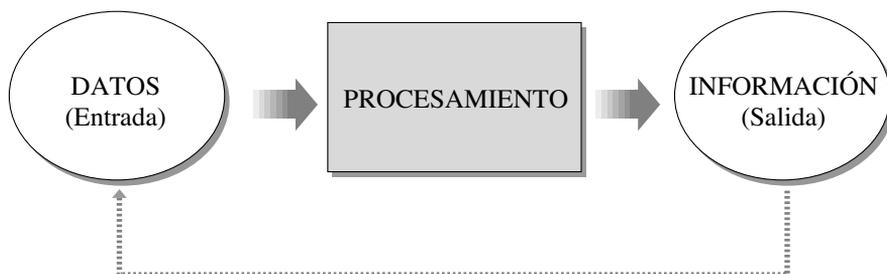


Fig. III.1 Sistema básico de información

Parece sencillo. Únicamente se trata de introducir unos datos, pulsar alguna tecla y obtendremos la información que deseábamos. Ahora bien, ¿qué datos vamos a introducir?, ¿cómo són?, ¿están bien *dispuestos* para el ordenador?, ¿cómo los va a entender la máquina?, ¿qué tecla, orden, comando o programa debemos accionar?, ¿es realmente Información lo que vamos a obtener o va a ser otro montón de datos más con los que no vamos a saber qué hacer? Esta situación es muy frecuente por diversos motivos, siendo el más común el de no cuidar el proceso de recogida, codificación y entrada de datos. (Normalmente, se dice que si a un ordenador le metes basura, basura sale). No es lo mismo disponer de datos que disponer de información³.

Partamos del supuesto de que nuestro estudio va a necesitar el diseño de un protocolo de recogida de datos. Supongamos además que para el tratamiento posterior de dichos datos decidimos optar por la ayuda de un ordenador.

Las líneas que siguen son una serie de consideraciones que serían deseables que toda investigador que esté en esta situación las tenga siempre presentes.

Recuerde que siempre es mejor *ganar* un poco de tiempo diseñando nuestro trabajo antes de encender la máquina y ponerse a meter datos de forma alocada. Las consecuencias se pagan después en horas de pantalla y teclado innecesarias. Tenemos que tener muy claro qué queremos conseguir, a dónde queremos llegar y después poner los medios para ello.

Vamos a introducir una mínima terminología que consideramos necesaria para poder continuar adelante.

Llamaremos a partir de ahora **Registro** a la colección de datos recogidos acerca de un individuo, entidad o cosa. Llamaremos **Campo** a cualquiera de estos datos recogidos (ej: *edad, sexo, estado civil, ...*). Los campos, por tanto, acogerán los **Valores** cada registro presenta. Por ejemplo, la edad, el peso, la talla, el sexo pueden ser campos que componen un registro, mientras que 36, 78, 1.60 y 2 pueden ser, respectivamente, los valores de un individuo (registro) para esos campos. En un entorno más puramente estadístico hablaríamos de OBSERVACIONES (*registros*) y VARIABLES (*campos*) de forma similar. Informáticamente hablando, todos los registros objeto de nuestro estudio deberán ser grabados en soporte magnético (disco) bajo la forma de un **Fichero**.

Lo primero y muy importante será, por tanto, definir con claridad qué objetivos perseguimos en esta fase de nuestro estudio. ¿Qué queremos hacer con los datos? ¿Cómo los queremos conservar? ¿Queremos mantener el Fichero de datos actualizado en el tiempo? ¿Vamos a realizar análisis estadísticos con los datos? ¿Qué tipo de análisis? ¿Queremos *cruzar* variables? ¿Cuáles?

En muchas ocasiones, hemos visto cómo usuarios que no tenían planteados unos mínimos objetivos con respecto a sus variables, ante la idea de cruzar variables, optan por "*cruzar todas con todas, a ver lo que sale...*". Y lo que sale es una montaña de hojas de papel llenas de tablas y más tablas de porcentajes y frecuencias. Esto impresiona mucho pero vale de bien poco. El análisis de estas tablas llevaría meses, si no años. Nuevamente, obtenemos a la salida del proceso únicamente más datos, en lugar de información. Recordemos que **la información es el dato utilizable**.

Una vez que tenemos claro cuáles son los objetivos que perseguimos en esta fase de la investigación (en definitiva, la información que queremos obtener), debemos seguir tres pasos, muy simples, pero que a menudo se ignoran:

- 1) Determinar exactamente cuál va a ser nuestra **unidad de registro**.
- 2) Diseñar la **estructura de registro**. Se trata de definir exhaustivamente los campos de que va a constar cada registro, elaborando lo que llamaremos un *diccionario de datos*.
- 3) Elegir el **soporte** adecuado para nuestros datos. Nos referimos con esto, al formato en que se van a grabar los datos en el ordenador. Este formato vendrá posibilitado por el programa de entrada de datos que utilizemos (editor de textos, gestor de base de datos, hoja de cálculo...).

Para todo ello, nos guiaremos siempre por el siguiente criterio:

Minimizar la cantidad de datos, Maximizar su calidad

2. Determinación de la unidad de registro

Se trata simplemente de elegir el sujeto u objeto sobre el cual vamos a recabar información y estructurar ésta en forma de campos. Supongamos que estamos realizando un estudio de calidad de aguas potables en una comunidad autónoma. Nuestro objetivo es establecer la existencia o no de diferencias entre los municipios y posteriormente entre las ciudades. Podríamos pensar que nuestra unidad de registro será por tanto el municipio. Sin embargo, nuestra fuente de datos indican que las tomas para los análisis se realizan siempre en determinados puntos de recogida para cada municipio: pozos, fuentes públicas o plantas depuradoras. Parece entonces que cada registro que tomemos contendrá los datos de análisis por cada punto de recogida de muestras. Sin embargo, sabemos también que para cada punto se repiten los análisis de agua cada cierto tiempo dentro de un año. La unidad de registro será, por tanto, cada análisis de aguas que se haga. Este registro contendrá, además de los datos propios del análisis, otros datos identificativos como el código del punto de toma, el código municipal y provincial y la fecha de recogida de la muestra.

Con este ejemplo, queremos indicar que no siempre el objeto de nuestro estudio de investigación coincidirá con la unidad de registro que nos veremos obligados a tomar. La **f fuente de datos** a la que podamos acceder determinará en gran medida esto.

Es aconsejable siempre pensar en nuestros datos en forma de matriz rectangular. Las entradas horizontales (*filas*) indicarían los registros u observaciones individuales,

mientras que las entradas verticales (*columnas*) señalan los campos o variables. (fig. 2)³. Un fichero sería el conjunto de valores para cada variable ó campo de cada individuo o registro, grabados en un disco magnético bajo un nombre.

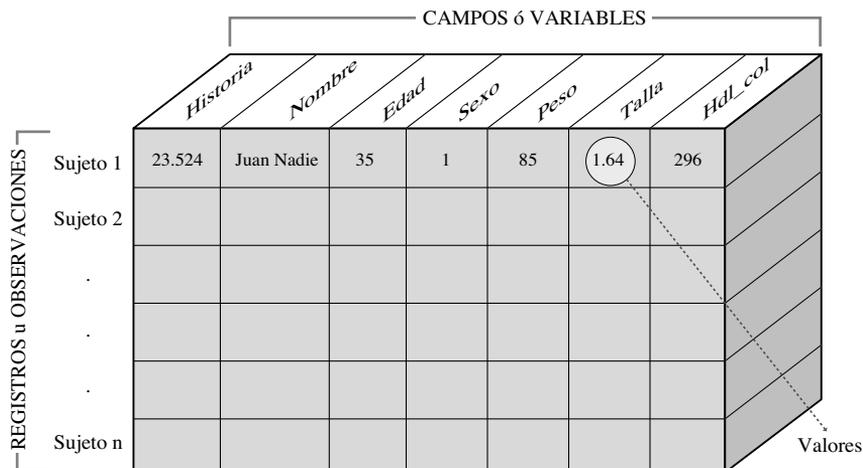


Fig. III.2 Sistema básico de información.

Si nosotros no somos capaces de organizar nuestros datos de forma que podamos representarlos matricialmente, el ordenador lo será aún menos. Esto no quiere decir que informáticamente hablando no se puedan manejar estructuras de datos diferentes como son las estructuras jerárquicas ó arbóreas. Lo que ocurre es que el uso de estas estructuras es demasiado complejo: se necesita no sólo un conocimiento más profundo de informática, sino también el dominio de sistemas de ordenadores y *software* más complejos.

Es probable que en alguna ocasión se nos presente una estructura de datos que a primera vista sea imposible de organizar matricialmente. El siguiente ejemplo puede ser muy esclarecedor.

Supongamos que estamos llevando a cabo un estudio en hogares: pensamos que el hábitat familiar influye en la estructura física de sus miembros (peso, talla,...). Parece que nuestra unidad de estudio debe ser, tanto, el hogar. De cada uno recogeremos una información básica: servicios, hábitos culinarios, zona, etc... De cada miembro de los convivan en el hogar también tenemos que recoger otro tipo de información: edad, sexo, peso, talla, etc...

Intentemos con esto construir una matriz rectangular. Cada fila de esta matriz dará información sobre un hogar distinto. Hay unos datos fijos que ya podemos colocar en cada columna:

Tabla III.1 Matriz rectangular de hogares (unidad de estudio) y características seleccionadas.

	Zona Viv.	Antigüedad	NºServic.	Ascensor	Edad Hijo_1	Edad Hijo_2	...?...
Hogar 1							
Hogar 2							
...							

Ya nos hemos encontrado con el primer problema: ¿Cuántas columnas tengo que reservar para la información de los miembros de cada hogar? ¿Cuántos son? Evidentemente cada familia tendrá un número diferente de ellos. Además pueden convivir otros familiares: primos, abuelos... Si intentamos representar matricialmente estos datos, nos saldría algo así:

Tabla III.2. Matriz dentada de hogares y características seleccionadas.

Hogar 1							
Hogar 2							
...							
...							
...							
...							
...							
Hogar N							

¿Qué es esto? Pues si lo de antes era una matriz rectangular, esto es lo que se llama una *matriz dentada*⁴. Una herramienta informática capaz de manejar estructuras jerárquicas ó arbóreas como esta del ejemplo nos vendría muy bien. Lamentablemente no están a nuestro alcance. Ya comentamos antes las limitaciones que nos suponen este tipo de herramientas. El *software* de uso más común y que está al alcance de cualquiera (vease epígrafe C) no maneja este tipo de estructuras. Tenemos que utilizar siempre estructuras matriciales. ¿Qué podemos hacer entonces?, ¿abandonar el estudio? Seguro que tiene que haber una solución.

Vamos a convertir este *fichero de datos* o matriz en dos. Es decir vamos a crear dos matrices, una contendrá exclusivamente los datos de cada hogar. Además identificaremos a cada hogar con un código (pueden ser un numeros consecutivos). Así este *fichero-matriz*, se configurará:

Tabla III.3. Matriz rectangular de hogares y características seleccionadas.

	Cód.Hogar	Zona Viv.	Antigüedad	NºServic.	Ascensor	Nº Miembros
Hogar 1						
Hogar 2						
...						

Ahora haremos lo mismo con los datos que deseamos recoger para cada individuo:

Tabla III.4. Matriz rectangular de individuos y características seleccionadas.

	Cód.Hogar	Parentesco	Edad	Sexo	Peso	Talla
Indiv. 1						
Indiv. 2						
...						

Esto ya está más claro. Ahora tenemos dos matrices perfectamente rectangulares. En la primera introduciremos los datos de los todos los hogares que visitemos (una fila por hogar). En la segunda irán los datos relativos a todos los individuos de todos los hogares.

Obsérvese que en esta segunda matriz también hemos añadido el código el hogar al que pertenece cada individuo. Así tenemos perfectamente identificados y relacionados a cada individuo con su hogar. El *software* que utilizemos se encargará de eso.

Se podría argumentar que podríamos utilizar como patrón esta segunda matriz, añadiéndole las columnas (variables-campos) que describen el hogar en que vive cada individuo. Esto puede ser válido y funcionar igualmente, pero no es óptimo. Tendríamos que repetir mucha información innecesariamente (*si tenemos a alguien que introduzca datos, que sea gratis y que no se proteste demasiado, podemos optar por esta solución*). Tenemos que recordar una vez más nuestro principio: maximizar la calidad, minimizar la cantidad.

3. Diseño de la Estructura de Registro

Una vez localizada y determinada la unidad de registro y que veamos nuestro datos en forma de matriz rectangular, procederemos a elegir y definir los campos que necesitaremos. Es decir, tenemos claro ya cuáles son las filas; se trata ahora de definir exhaustivamente cuántas y cómo serán las columnas.

Respecto al número de campos a elegir de la fuente de información, hay dos posturas encontradas: Una primera nos aconseja introducir el mayor número posible de datos para cada registro, independientemente de que muchos de ellos no los necesitemos para nuestro estudio. Esta postura, un tanto conservadora, viene de la idea de que *quizás* después podamos necesitar algún dato que no habíamos previsto, o bien por la razón de que si una vez terminado nuestro estudio queremos continuar ampliándolo con nuevos datos, éstos se hallan ya recogidos e introducidos en la máquina.

Una segunda postura, más económica y racional, nos dice que si tenemos bien definida nuestra metodología de trabajo, elegiremos unicamente el subconjunto de variables estrictamente necesarias. Esta segunda postura, además de resultarnos más cómoda a la hora de introducción de datos, resulta mucho más económica en tiempo y espacio de máquina necesarios. Además nos proporcionará un mayor control de calidad sobre nuestros datos.

Una vez que designemos cuáles van ser los campos que compondrán nuestro registro, procederemos a describir para cada uno una serie de características. Esta definición es lo que denominamos **construcción de un Diccionario de Datos**⁵ y es muy importante, ya que determinará en gran medida el medio (*software*) por el cual serán introducidos los datos en el ordenador. Por otro lado, el mero hecho de sentarse y pensar un poco sobre cómo son nuestros datos hará que aparezcan en este momento todas las dudas y dificultades que tengamos; después ya sería demasiado tarde. Una razón más: la comunicación con un informático siempre debe hacerse en estos términos.

Por tanto, se trata de explicitar para cada campo el mayor número posible de características, como pueden ser:

- a) Nombre del campo.
 - b) Tipo de dato (campo).
 - c) Rango: valor mínimo y valor máximo previsible.
 - d) Anchura máxima: número de caracteres del valor máximo previsto.
 - e) Valor faltante (*missing value*)
 - f) Campo codificado o valor directo
 - g) Campo derivado del valor de otros campos o no derivado.
- a) **Nombre del campo:** comenzaremos por bautizar al dato que deseamos recoger, pensando ya en la máquina y en el programa que vayamos a utilizar para procesar los datos. La gran mayoría de los programas identificarán los campos o variables con nombres que no pueden exceder de 8 caracteres. Si uno de

nuestro campos es la edad, lo tenemos fácil: le llamaremos *edad*. Pero si se trata de la presión arterial sistólica, lo tenemos más complicado; podemos elegir como nombre para este campo: *PAD*. Procederemos así con todos nuestro campos ó variables.

- b) **Tipo de campo:** Se trata de definir cómo van a ser los valores que puede contener cada campo. Así la variable edad es una *variable numérica* ya que los posibles valores que pueda contener son dígitos y además expresan una cantidad. Hablaremos de *variables alfanuméricas* cuando los posibles valores que pueden tomar son tanto dígitos como letras del alfabeto. Las que más nos interesan de éste tipo son las variables codificadas. Así, el sexo de un individuo podemos codificarlo como 'H' y 'M', o bien como '1' y '2' para hombres y mujeres respectivamente. Atención a esto, porque algunos programas estadísticos no admiten caracteres alfabéticos para cierto tipo de análisis.

Por tanto, esto es muy importante definirlo previamente, de cara al programa de tratamiento que posteriormente utilizemos con el ordenador, ya que le deberemos explicitar si se trata de un tipo u otro de variable (numérica o alfanumérica). Posteriormente volveremos sobre el caso de las codificaciones.

Además de campos de tipo numérico o alfanumérico, encontramos con frecuencia *campos especiales* que no encajan en uno u otro tipo. Nos estamos refiriendo, en concreto, a las fechas y a las horas. En efecto, son muchos los estudios en investigación sanitaria que necesitan jugar con fechas, como son los estudios de supervivencia, en los que es necesario a menudo elaborar cálculos a partir de dos o más fechas. Los estudios apoyados en teoría de colas o simulación necesitan a menudo trabajar con horas de entrada y salida del sistema, reduciéndolas despues a la diferencia en minutos. Como después veremos, la existencia de estas variables en nuestro estudio determinará en gran medida la elección de determinados programas de ordenador para trabajar con los datos.

Por último, hay otro tipo de campos denominados *campos lógicos*, llamados así porque únicamente admiten dos valores sobre el objeto que representan: verdadero o falso. Este tipo de campos podían ser considerados como campos alfanumericos codificados (valor '0' ó '1') pero la existencia de gestores de bases de datos como dBASE⁶ que manipulan este tipo de campos con gran seguridad hace aconsejable su uso. Por cierto, el sexo..., ¿es un campo lógico? No, ya que un campo lógico se reserva para aquellos valores de los que se admite un estado de verdad o falsedad. Un libro, lo tenemos prestado o no. Una persona ha sido operada o no. Pero no podemos decir si tiene sexo o no. Puede ser hombre, mujer o desconocido (3 valores).

- c) **Rango:** Se trata únicamente de establecer cual puede ser el valor mínimo y el valor máximo que puede soportar el objeto a que se refiere este campo. En el caso de campos codificados con letras, se indicará la menor y mayor letra en orden alfabético. Si está previsto que el campo se quede en blanco, este será el valor mínimo.
- d) **Longitud o anchura del campo o variable:** Debemos preveer que número de caracteres o dígitos van a tener los valores de cada una de las variables que estemos manejando. En el caso, de las variables numéricas, anotaremos también el número de decimales posibles, si es que los tienen.
- e) **Valores Faltantes** (missing values): Previamente a la entrada de los datos en el ordenador debemos pensar también en los posibles valores faltantes para cada variables; me refiero a que el programa informático debe de asignar algún valor para reconocer aquellos campos (variables, preguntas) que nos ha sido imposible recoger, es decir aquellas variables cuyo valor se desconoce por algún motivo. Para esto, *asignaremos un valor que quede fuera del rango posible de valores* de la variable en cuestión. Así, si encontramos un cuestionario en el que no se ha cumplimentado una pregunta podemos asignar el valor 'X'. Para esto hay ya convenciones que se suelen utilizar. Así, con frecuencia encontramos el '9' como valor faltante más utilizado. En el ejemplo anterior sobre el sexo, si tenemos a éste codificado como '1' y '2', en caso de que lo desconozcamos, podemos optar por el valor '9'
- f) **Campos codificados:** Anotaremos también en nuestro diccionario si el campo que estamos definiendo es un campo codificado o no. Si la codificación es compleja es conveniente acompañar siempre una hoja de codificaciones. En lo que respecta a los códigos es el momento de hacer dos advertencias: una respecto a las variables cuantitativas que se categorizan, codificándolas antes de entrar al ordenador y otra respecto a la utilización de códigos en ciertas preguntas en cuestionarios.

En primer lugar, es muy importante el no codificar previamente a la entrada al ordenador aquellas variables numéricas (cuantitativas) que las deseemos después categorizar (codificar), ya que siempre y en cualquier momento podemos dar la orden oportuna al programa de tratamiento para que haga este trabajo por nosotros. Por ejemplo, si una de nuestras variables es la edad, es mejor recoger el dato directamente en lugar de categorizarlo en grupos de edad. Así podemos siempre jugar con distintos intervalos de edad, mientras que si introducimos la edad directamente codificada nunca podremos volver atrás para cambiar los intervalos en base a los valores reales que ya no tenemos. Además siempre tendremos a nuestro alcance dos tipos de análisis para esa variable.

Cualquier programa que se precie debe dar la oportunidad de recodificar los valores de las variables según nuestros deseos.

El caso de las preguntas abiertas y preguntas multirespuesta: Puede ocurrir que en el diseño de un cuestionario, decidamos utilizar preguntas de tipo abierto así como preguntas multirespuesta en la que ofrecemos un abanico de posibles respuestas, pudiéndose marcar una o más de una.

En cuanto a las primeras, preguntas abiertas, no cabe en modo alguno su "informatización" directa, aunque ya hay algún programa de *análisis textual*⁷ que es capaz de analizar expresiones y categorizarlas. A falta de esto, lo que si se puede hacer es realizar un análisis lógico de las respuestas hechas y agruparlas en posibles categorías que pueden ser susceptibles de categorización y por tanto de tratamiento con el ordenador. Se trataría de hacer un inventario lo más extenso posible de todas las posibles respuestas. Claro está que esto es operativo sólo si el número de cuestionarios es reducido. (Piénsese en la gran variedad de respuestas que pueden ser recogidas con una muestra de 1500 individuos a una pregunta tan abierta como *¿qué piensa sobre nuestro sistema sanitario?*).

En cuanto a las preguntas tipo multirespuesta, realmente no son una sólo variable sino tantas como respuestas se hallan dado. El problema está en que siempre, como he dicho antes, el programa informático debe de conocer con cuántas variables se va a trabajar. En efecto, ante una pregunta de este tipo en un cuestionario, unas personas pueden elegir tres opciones, otras dos, otras ninguna, y otras, muy concienzudas, las marcan casi todas. Con el siguiente ejemplo lo veremos mejor; en un cuestionario se ha realizado la siguiente pregunta:

¿Qué tipo de formación cree necesaria para abordar un estudio en salud materno-infantil?

(Multirespuesta)

- 1.- Psicología
- 2.- Psiquiatría
- 3.- Pediatría
- 4.- Sociología
- 5.- Tocoginecología
- 6.- Antropología
- 7.- Medicina General
- 8.- Puericultura
- 9.- Ética y Moral
- 10.- Pedagogía

Nótese que el autor de este cuestionario, al dar la posibilidad de multirespuesta, abre la posibilidad de un número de combinaciones extremadamente elevado. No sólo una persona puede elegir un número variado de respuestas, sino que además no puede establecer orden de prioridad ninguno.

Aunque hay varias soluciones, ante un problema así y ya *a posteriori* con el cuestionario recogido, la mejor solución es convertir lo que aparentemente se suponía que era una sola pregunta en 11 preguntas a las que se ha podido optar entre el SI o el NO. Esto significa pasar de tener una variable a 11 de repente. Es decir, podríamos codificar los valores de cada posible respuesta como '1' cuando se ha elegido y '0' cuando no se ha elegido.

También podríamos haber optado por formular la misma pregunta en otros términos:

"De los siguientes, señale en orden de preferencia, los tres formaciones que le parezcan más importantes para abordar un problema en salud materno-infantil"

De esta forma, el número de variables que tenemos que manejar es de 3, en lugar de 11, y el rango de valores posibles para cada una de ellas va de 1 a 11, en lugar de 0 a 1.

Otra posible solución, si queremos proteger la unidad de la pregunta, es codificar las respuestas de forma directa. Así si se han elegido las opciones 2, 5 y 8, el valor de este campo para este registro será '258'. El problema es que hay que definir la anchura máxima posible, para este campo será de 11 caracteres. Amén de que el número resultante de combinaciones posibles hace tremendamente complejo el análisis posterior de esta variable.

No es el propósito de esta líneas, juzgar cuál de los métodos es más correcto para estas situaciones. El análisis que se deriva de una u otra es diferente así como posiblemente la información que podamos extraer. Nuestro propósito es sólo informar de lo distinto que puede ser el tratamiento informático en una u otra forma, y que el investigador debe de conocer previamente este problema antes de diseñar el cuestionario.

- g) **Campos derivados:** Por último, planificar siempre tiene sus ventajas. Hay campos que podemos extraer a partir del valor de otros campos. Por tanto, no debemos gastar tiempo en recogerlos e introducirlos en el ordenador. Supongamos que hemos recogido el peso y la talla de una serie de individuos; necesitamos además tener el índice de masa corporal, (o *Índice de Quetelet*), que

como sabemos es la siguiente relación **Peso/Talla**². Debemos saber que cualquier programa que se precie debe poder ofrecer la posibilidad de "elaborar" estos nuevos campos o variables a partir de otros, ahorrándonos un considerable trabajo.

Es aconsejable seguir todos estos pasos y plasmar cada definición de variables en lo que podríamos llamar *diccionario de datos*. Si hemos seguido este proceso para cada una de las variables, estamos ya en disposición de introducir los datos en el ordenador. A continuación se ofrece un ejemplo:

Unidad de registro: cada intervención quirúrgica realizada en el el Hospital Comarcal Tal desde el 1 de Enero al 30 Junio de 1.991.

Tabla. III.5. Diseño de registro de campos derivados.

	Campo	Tipo	Codif.	V.Min	V.Max	Long	Missing	Derivado
1	Numero	N		1	10000	5		
2	Quirofano	A	Si	1	10	2	99	
3	Tipo_int	A	Si	001	666	5	999	
4	Origen	A	Si	P	U	1	X	
5	Fecha	F		1-1-91	30-6-91	8	Blanco	
	<i>Hora_ini</i>							
6	hora_i	N		0	23	2	99	
7	minuto_i	N		0	59	2	99	
	<i>Hora_fin</i>							
8	hora_f	N		0	23	2	99	
9	minuto_f	N		0	59	2	99	
10	Duracion	N		1	1440	4	0	SI ^(*)
11	Ingreso	L		F	V	1		
12	Sexo	N	Si	1	2	1	9	
13	Edad	N		1	99	2	0	
14	N_Historia	A		1	1000000	7	Blanco	

(*) La duración de la intervención la estimaremos restando la hora de inicio de la hora de finalización.

4. Elección del formato de datos

Una vez diseñado el registro, tenemos claro los campos a utilizar y las características de cada uno. Ahora, y no antes, es cuando llega el momento de utilizar el ordenador para la introducción y tratamiento de los datos.

Para lo primero, utilizaremos un determinado *software* de entrada de datos. Existen varias opciones, aunque recomendamos que, para un trabajo de investigación, los datos se introduzcan mediante el mismo programa que va a servir de análisis y tratamiento. Cada opción tiene sus propias ventajas e inconvenientes, pero sea cual sea el programa elegido, debemos asegurarnos de que tenga **comunicación con otros programas**, por si en un momento determinado deseamos transferir nuestro fichero de datos a otro programa. Los tres formatos más comunes de entrada de datos son:

- 1) En formato ASCII, mediante un editor de texto: el editor del Sistema Operativo EDIT, un editor específico como Qedit, etc... Hay que decir que casi todos los programas incorporan su propio editor. Así, SPSS/PC proporciona el editor REVIEW, SAS incorpora su propio sistema de ventanas de edición...
- 2) En formato dBASE, mediante este mismo programa. Es quizás la forma más utilizada hoy.
- 3) En formato Hoja de Cálculo, mediante Lotus o Quattro-Pro.

Debemos saber que casi todos los programas de análisis estadístico permiten tratar los datos grabados previamente en formato ASCII. Esto haría pensar que el mejor método para introducir los datos es hacerlo sobre un editor de texto en este formato. Sin embargo, un editor de texto no nos proporciona ningún control sobre los datos que estamos tecleando, mientras que LOTUS o el dBASE sí lo hacen y de forma mucho más efectiva.

Además debemos saber también que estos dos programas no sólo puede transferir los datos en formato ASCII sino que además tienen comunicación directa entre ellos. Así, podemos llamar directamente desde SPSS, desde EPIINFO o desde SAS un fichero creado con dBASE.

Escribir directamente los datos sobre un editor es recomendable en las siguientes situaciones:

- 1) La cantidad de registros es escasa (menos de 200)
- 2) El número de campos es pequeño. En total, deben de sumar menos de 80 caracteres (la anchura de la pantalla). Cada registro debe ocupar una línea de pantalla.
- 3) Si estamos muy seguros de la calidad de los datos y de nuestra habilidad con el teclado para no cometer errores.
- 4) El fichero no necesita mantenerse en el futuro, es decir, una vez introducida a totalidad de registros, no añadiremos ni borraremos ni modificaremos registros.

La elección de una Hoja de Cálculo es recomendable cuando:

- 1) Deseamos realizar transformaciones muy a menudo con los datos para averiguar indicadores tabulados en vertical y horizontal (tasas)
- 2) El número de columnas y filas no debe ser muy elevado: la hoja de cálculo se enlentece mucho y ocupa demasiada memoria.
- 3) Vamos a trabajar con horas y fechas de forma continua. Es el caso de estudios de teoría de colas y simulaciones. En general, los estudios derivados de Técnicas de Administración aconsejan la utilización de la Hoja de Cálculo.
- 4) Deseamos realizar gráficos sobre nuestros datos.

La elección de un gestor de bases de datos como dBASE se aconseja cuando:

- 1) El número de registros es muy elevado
- 2) El número de campos es alto. dBASE III admite hasta 128, mientras su hermano mayor, dBASE IV, admite hasta 1024.
- 3) Deseamos establecer controles de calidad sobre nuestros datos: depurar errores.
- 4) Deseamos mantener nuestro fichero vivo: altas, bajas, modificaciones a lo largo del tiempo.
- 5) Deseamos hacer consultas relacionales; ej: "*¿qué mujeres mayores de 35 años fueron operadas de cesárea en el quirófano 'B' entre el 1 de Marzo y el 31 de Marzo del 91?*"
- 6) Cuando deseamos enviar de forma selectiva sólo algunos campos y registros (dependiendo de determinadas condiciones) para su tratamiento con algún programa de análisis estadístico ó epidemiológico.
- 7) Cuando deseamos mantener tablas de códigos, adjuntas a la base de datos principal.

En general, para elegir el programa de tratamiento e introducción de datos, responderemos a estas cuestiones:

- Número de registros previsibles
- Número de campos

-
- ¿Vamos a operar con campos especiales: fechas, horas...?
 - ¿Vamos a realizar operaciones "intra-registro" (tratamiento vertical u horizontal)?
 - ¿Necesitamos mantenimiento de códigos?
 - ¿Necesitamos depurar los datos de posibles errores?
 - ¿Van a sufrir transformaciones algunos campos?
 - ¿Deseamos mantener el fichero vivo?
 - ¿Cómo va a ser la explotación de nuestros datos?:
 - ¿Qué tipo de análisis vamos a realizar?
 - ¿Qué programa hará este análisis?
 - ¿Cómo necesita este programa que estén los datos (formato)?

C. SOFTWARE DISPONIBLE

A. Calzas Urrutia

En este epígrafe daremos un brevísimo repaso al *software* (conjunto de programas) existente, intentando proporcionar más criterios para elegir entre un programa u otro a la hora de manejar nuestros datos en el ordenador.

Para los objetivos que perseguimos, podemos clasificar el *software* de utilidades para usuarios en cuatro grandes apartados:

- Procesadores y/o Editores de Texto.
- Hojas de Cálculo y Gráficos.
- Gestores de Bases de Datos.
- Paquetes Estadísticos.

Estas son las herramientas básicas con las que debemos contar en cualquier momento si queremos realizar un tratamiento informatizado de nuestro conjunto de datos, además de un mínimo conocimiento del sistema operativo. Evidentemente no siempre es necesario utilizar todas en un mismo proyecto de investigación. La elección de un programa u otro vendrá determinada por unos criterios que abordaremos más tarde.

Pero antes creo necesario dar una breve descripción de lo que es cada una de aquellas utilidades enumeradas, intentando sobre todo explicar qué se puede hacer con cada uno de ellos, para qué nos pueden servir.

1. Procesadores o editores de texto

Enumerar las funciones de un procesador de textos sería demasiado largo y costoso. Inicialmente diremos que podemos utilizar un procesador de texto para un doble objetivo: para escribir y editar cualquier tipo de documentos y para introducir datos "en crudo".

- 1) Como herramienta de entrada de datos es también muy útil aunque quizás demasiado engorroso, ya que el procesador de textos no tiene por qué distinguir si lo que estamos metiendo son datos pertenecientes a variables o una lista de la com-

pra, por decir cualquier cosa. Es decir, disponemos únicamente de una página (pantalla) en blanco, en la que podemos introducir los datos (variables) pertenecientes a cada caso en cada línea de texto. No tiene control de errores; así si se nos olvida introducir un valor, este error pasará desapercibido hasta que empiece a dar problemas cuando estemos analizando los datos con un paquete estadístico.

- 2) Como editor de documentos, diremos que un procesador de textos simula perfectamente a cualquier máquina de escribir, mejorándola. Por lo pronto, no hay necesidad de ejecutar el retorno de carro cada vez que lleguemos a un final de línea. El procesador de textos se encarga de forma automática de continuar al principio de la línea siguiente del párrafo, sin cortar la palabra que estábamos escribiendo y, además, con la obtención del margen derecho justificado (como se está viendo en este mismo texto).
- 3) Además de fijar todas las características de presentación de nuestro documento (como márgenes, tabuladores, espaciado entre líneas ...) podemos jugar con cierta comodidad con todos los elementos de un texto: caracteres, palabras, líneas, párrafos, páginas... Una de las mayores ventajas es la de recomponer un párrafo escrito según unas determinadas características de acuerdo a otras que fijemos posteriormente sin necesidad de volver a escribirlo o hacerlo de forma manual.
- 4) Además, uno de los elementos más importantes es el bloque de texto. Si marcamos un bloque de texto ya escrito, podemos copiarlo en cualquier otro lugar del documento, moverlo o trasladarlo de sitio, borrarlo de un solo golpe de tecla, e incluso transportarlo a otro documento diferente, aprovechándolo sin tener que volver a escribirlo de nuevo.
- 5) Podemos también buscar de forma automática cualquier palabra o palabras a lo largo del documento y opcionalmente sustituirla también de forma automática por cualquier otra secuencia de caracteres. Esto sin mencionar a las utilidades de Ortografía y diccionario de Sinónimos que ofrecen los procesadores de texto.

En definitiva, la gran ventaja del proceso de textos es la posibilidad de abstraernos a la tiranía del papel y lápiz: cualquier cosa puede ser modificada en cualquier momento.

En el mercado existen una gran variedad de procesadores de texto, siendo los más populares: WordPerfect, Microsoft Word, WordStar...⁸

Además hay editores de texto, que no tienen la función de proceso de textos y cuya finalidad es servir únicamente como herramienta de entrada de datos y escritura de programas. Quizás uno de los más adecuados sea el QEdit por la velocidad con

que se trabaja con él y lo sencillo de su manejo. Este programa además es de los conocidos como *shareware*: no se necesita licencia, su distribución es libre, aunque el usuario está moralmente obligado a aportar una pequeña cantidad (casi simbólica) por su utilización: a cambio recibirá los manuales y las versiones actualizadas. En cualquier caso, el sistema operativo MS-DOS ofrece también su propio editor llamado EDIT.

2. Hojas de cálculo

Una Hoja de Cálculo (u hoja electrónica) es un programa que permite disponer datos en filas y columnas y poder realizar con ellos un elevado número de **funciones** así como cualquier tipo de operación aritmética con ellos (**fórmulas**). Además de ello, podemos obtener **gráficos** inmediatos de los datos contenidos en la Hoja. Otra de las funciones de la Hoja de Cálculo es su posibilidad de actuar como una **base de datos** rudimentaria.

No todas las hojas electrónicas existentes en el mercado disponen de todas las posibilidades antes descritas. Las más completas son Quattro-Pro⁹, Lotus 1-2-3, Excell,... Cualquiera de las mencionadas integra las tres funciones descritas (operaciones, base de datos y gráficos), tienen una gran facilidad de comunicación con otros programas (los datos pueden ser trasladados) y además se hallan en castellano; todos los comandos están traducidos a nuestro idioma así como el programa interno de ayudas que lleva.

¿Qué nos ofrece a nosotros la Hoja de Cálculo? El gestor y/o administrador sanitario lo verá enseguida muy claro: puede seguir utilizando la Hoja de Cálculo en su sentido más primario: como herramienta de gestión.

Pero, en realidad, podemos utilizar la Hoja de Cálculo para cualquier tarea que suponga el manejo de datos tabulados: cálculo de tasas y su representación gráfica, simulación de demanda, control de gastos, gestión de almacén ... En concreto se ha utilizado muchísimo dentro del entorno de la Epidemiología en elaboración de indicadores, tasas y representación gráfica, fundamentalmente.

Para dar una idea, describiremos muy brevemente las funciones de una Hoja de Cálculo potente, como pueda ser Lotus o Quattro-Pro. Veremos cómo todas están enfocadas hacia el análisis y toma de decisiones.

2.1. Análisis y tratamiento de datos

Se trata de aplicar la memoria y la velocidad de los ordenadores a problemas que de otra forma deberían resolverse manualmente con papel, lápiz y, en el mejor de los casos, con una calculadora. Podemos disponer de una inmensa hoja de 8.192 filas por 256 columnas y retener en ella todo tipo de tablas de números relacionados con complejas fórmulas.

Inmediatamente podríamos observar una de las funciones más potentes de la Hoja: su posibilidad de **recálculo**. En efecto en cualquier momento, podemos modificar uno o más datos de entrada y observaríamos cómo automáticamente se recalculan todos aquellos resultados que se ven afectados por el cambio realizado. Esto nos lleva a una interesante función para la toma de decisiones. Se trata del "qué ocurriría si ..."; En efecto podemos jugar cambiando las entradas y obteniendo de forma inmediata el efecto de aquellos cambios, sin necesidad de espera.

2.2. Gestión de la Información.

Los usuarios de la Hoja de Cálculo pueden desear organizar la información contenida en la hoja de trabajo. Para ello, el programa puede hacer ordenar el contenido alfabética o numéricamente en orden ascendente o descendente. Podemos también ordenar que localice y extraiga datos de una hoja utilizando hasta 32 criterios de búsqueda o selección.

2.3. Representación Gráfica.

En cualquier momento podemos elegir un rango de valores nuestra hoja de trabajo y ordenar al programa que los represente gráficamente ya sea en gráficos de barras, sectores, líneas, coordenadas XY...

3. Gestores de bases de datos

Hasta ahora hemos venido utilizando una terminología bastante más cercana a la Estadística: hablabamos de Observaciones y Variables. En un entorno más informático, como es el de las Bases de Datos, estos conceptos se corresponden con Registros y Campos.

Una base de datos es un conjunto de información útil organizada de una manera específica; esta organización puede responder a muy diversas formas. Las más conocidas son las bases de datos jerarquizadas y las bases de datos relacionales. Nos referiremos a éstas últimas puesto que su modelo se está imponiendo actualmente.

Una base de datos relacional organiza los datos en forma de tablas con filas y columnas. Cada fila es un registro (un caso) y cada columna es un campo (variable). Con ayuda del Algebra de Boole, establece relaciones entre ambas entradas a la tabla de forma que la consulta a este tipo de bases se hace mucho más potente y efectiva.

A cada registro de un fichero de base de datos se le asigna automáticamente un número de registro. Los campos se identificarán por un nombre o etiqueta. En el siguiente ejemplo se ve mucho más claro:

Tabla. III.6. Registro de fichero de base de datos relacional.

NOMBRE	EDAD	SEXO	E_CIVIL	TABACO	ALCOHOL	...
Perez y Perez, A.	56	V	S	3	1	
Sanchez y Sanchez, P.	34	M	D	1	2	
Ruiz Ruiz, M.	24	V	C	3	5	
Porras y Porras, G.	39	V	C	0	1	
...						

Como se ve, se trata de una sencilla tabla de doble entrada: cada fila es un registro y cada columna es un campo al que se le ha dado un nombre identificativo.

En definitiva, a nosotros, como usuarios, lo que nos importa es la potencia de utilización de una base de datos: sus funciones más importantes residen la capacidad de localizar y acceder a cualquier elemento de información en múltiples modos (acceso directo a una ficha, búsqueda de individuos que cumplen determinadas condiciones, ordenación del fichero de datos según uno o más criterios ...)

Además nos puede servir como un gran depurador de errores. Una relación imposible es que un individuo sea hombre y tenga cancer de útero. Pero en la entrada de datos se producen muchos errores como estos, sobre todo cuando ambos campos están codificados. Podemos entocer, requerir del Gestor de Base de Datos, mediante una simple orden que nos extraiga de todos los registros aquellos cuyo campo Sexo sea igual a 'H' y cuyo campo Cáncer sea igual al código de localización de cáncer de útero. El Gestor de Base de Datos nos extre rapidismente los registros que cumplen estas condiciones, si los hay, con lo cual estamos en condiciones de modificarlos oportunamente.

4. Paquetes estadísticos

El auge de la informática y concretamente la accesibilidad que existe actualmente para utilizar determinados paquetes estadísticos ha popularizado el uso de la estadística como herramienta para el proceso de datos. Sin embargo, a pesar de considerar esta popularidad como beneficiosa, se están presentando una serie de problemas que pueden ser más o menos graves.

Fundamentalmente el problema que se plantea reside en el hecho de que la utilización de paquetes estadísticos obliga a una formulación más estricta y más clara en la definición y manejo de datos (vease Epígrafe B).

Además en ocasiones, hemos detectado un desconocimiento de las bases más elementales de los métodos que se utilizan, aplicándose erróneamente las técnicas que

se utilizan o bien utilizándose determinados métodos sólo porque el paquete disponible los contiene cuando el método idóneo sería otro, disponible en otro paquete.

Está claro, además, que para aplicar un programa estadístico o de epidemiología hay que saber algo de Estadística y/o Epidemiología. Es evidente la conveniencia de consultar a un especialista en estas materias antes de abordar un estudio que implique un diseño metodológico que no se conozca suficientemente. El profesional podrá aconsejar sobre el método estadístico más apropiado y las posibilidades que ofrece el paquete que se vaya a utilizar.

El *software* estadístico que existe actualmente en el mercado cubre un amplio abanico que va desde la utilización de las hojas de cálculo (ya comentadas) hasta paquetes muy sofisticados. Dejando a un lado las hojas de cálculo, dentro de los paquetes con una clara orientación al análisis de datos se pueden distinguir dos categorías:

1) De propósito general:

Incluyen un conjunto de métodos básicos más o menos estándar. De cara al usuario, podríamos dividirlos a su vez en interactivos y en los basados en un lenguaje de comandos u órdenes. Los interactivos están concebidos de forma que el mismo programa va guiando al usuario a través de un sistema de menús. Cabe mencionar entre estos a SIGMA, MICROSTAT y STATGRAPHICS.

El segundo subgrupo (basados en lenguaje de comandos) requiere que el usuario conozca un lenguaje de comandos propio, tecleando las instrucciones directamente. Por esta misma característica, estos paquetes son mucho más flexibles y potentes que los anteriores. A través de módulos opcionales cubren muchísimos procedimientos estadísticos. Entre ellos, cabe mencionar el SPSS¹⁰, el BMDP, SAS, Además estos programas, nos permiten escribir nuestros propios procedimientos en forma de subprogramas. Cabe mencionar en este grupo programas como GLIM¹¹, para el estudio de modelos lineales utilizando el método de máxima verosimilitud, S-PLUS¹², para análisis gráfico de datos, EPIINFO¹³ y EPILOG¹⁴, para análisis epidemiológicos.

2) De propósito específico:

Son programas para cálculos estadísticos específicos. En este grupo se sitúan programas como SPAD, para el estudio de análisis de correspondencias, LOGRESS, para regresión logística, SPAD-T, para análisis textuales, OSIRIS, para el tratamiento de cuestionarios, y un largo etcétera.

REFERENCIAS

- 1.- Prieto A, Lloris A, Torres JC. *Introducción a la Informática*. Madrid: Ed McGraw-Hill, 1989.
- 2.- Guilera L. *Introducción a la Informática*. Barcelona: Ed P.P.U., 1985
- 3.- Aparicio F. *Tratamiento informático de encuestas*. Madrid: Ed. Ra-ma, 1991
- 4.- Schinkel M, Kasper J. *Programming in Clipper 5*. New York: Addison Wesley, 1992
- 5.- Pressman R. *Ingeniería del Software. Un enfoque práctico*. Madrid: Ed. McGraw-Hill, 1988.
- 6.- Asthon-Tate. *Cómo utilizar dBase IV*. Madrid: Ed Borland International Inc., 1992.
- 7.- CISIA. *SPAD-T. Manual del usuario*. Madrid: Ed CISIA, 1992
- 8.- Cueva E. *WordPerfect 5.1*. Madrid: Ed Anaya, 1993.
- 9.- Borland. *Quattro Pro. Guía de actualización*. Madrid: Ed Borland International Inc., 1992.
- 10.- Etxeberría J, Joaristi L, Lizasoain L. *Programación y análisis estadísticos básicos con SPSS/PC+*. Madrid: Ed Paraninfo, 1990.
- 11.- Healy M. *GLIM: an introduction*. Oxford: Clarendon Press, 1988.
- 12.- StatSci. *S-Plus for DOS; User's Manual*. Seattle: Statistical Sciences, Inc., 1991.
- 13.- Dean A. *Manual de EPIINFO: Epidemiología con ordenadores*. Sevilla: Consejería de Salud. Junta de Andalucía, 1991.
- 14.- EPILOG. *Reference Manual*. Pasadena, CA: Epicenter Software, 1990

PARTE IV
ESCRITURA Y LECTURA CRÍTICA
DE ARTÍCULOS CIENTÍFICOS

Manuel A. Bobenrieth Astete
EDITOR

A. FORMAS DE COMUNICACIÓN CIENTÍFICA

M. A. Bobenrieth Astete

1. Tipos de publicaciones científicas

Hoy en día hay varias maneras de publicar información proveniente de investigaciones científicas realizadas en el campo de la salud. Esto se manifiesta en diversos tipos de publicaciones científicas.¹ De entre ellas se destacan:

- a. Revistas. Estas constituyen la fuente de información más numerosa de aparición en intervalos o períodos de tiempo regular definidos. Esta característica constituye la razón por la cual se las denomina «publicaciones periódicas». La publicación más frecuente es la mensual; luego la siguen la bimestral y la trimestral.

Las revistas se pueden clasificar de acuerdo a varios criterios no excluyentes:

1. Revistas de profesiones (médicos, enfermeras, odontólogos, farmacéuticos, etc.).
2. Revistas de especialidades de la profesión médica (urología, pediatría, cirugía, medicina, obstetricia, etc.).
3. Revistas de modalidades de atención (atención primaria, atención de urgencia, atención intensiva, atención de crónicos, etc.).
4. Revistas de investigación.
5. Revistas de enfermedades (cáncer, SIDA, etc.).
6. Revistas de literatura primaria.
7. Revistas de literatura secundaria, habitualmente resúmenes y/o revisiones.

La literatura biomédica es un quehacer muy complejo.² Los esquemas de clasificación son, en el mejor de los casos, simplificaciones artificiales. No debe sorprendernos, en consecuencia, que algunas revistas desafían cualquier clasificación estricta porque a menudo cumplen varios propósitos simultáneamente. De hecho algunas secciones de una revista médica pueden considerarse como parte de la literatura primaria, es decir información proveniente de estudios originales inéditos. Otras secciones corresponden a literatura secundaria o «infor-

mación sobre información». Igualmente encontramos un artículo de investigación original con metodología sofisticada al lado del informe descriptivo de un caso clínico novedoso. Ambos informes pueden aparecer junto a otro sobre la revisión de un tema de actualidad.

- b. *Libros*. Los libros varían en forma increíble en su origen, contenido, propósito, forma y disponibilidad. Así se los suele denominar libros de texto, libros de referencia, manuales, compendios, enciclopedias, series continuadas, suplementos con hojas sueltas, diccionarios, glosarios, guías, directorios, etc..

El libro tiene como característica la de ser una publicación independiente de sus autores. En otras palabras el autor de un libro, o aún el editor de una obra colectiva es -en sentido bibliográfico- un agente independiente. En efecto, goza de mayor libertad que el autor de un artículo científico original cuyo manuscrito se somete a una revisión por arbitros del consejo editorial de la revista. Este tipo de «arbitraje por pares» no sucede habitualmente con los libros ni con las monografías.

Los libros tienden a ser hoy en día el producto del esfuerzo de múltiples autores a lo largo de los diversos capítulos. Este hecho favorece la especialización de los temas a expensas, muchas veces, de una mejor cohesión y unidad de esa variedad de contribuciones.

Los libros presentan la ventaja de compilar información relevante en un mismo volumen. Presentan, sin embargo la desventaja de una más rápida desactualización de la información. De hecho a partir del momento de la decisión de escribir un libro, hasta la producción de la obra impresa y el inicio de su distribución, habrán pasado de 2 a 4 años.

- c. *Monografías*. Las monografías son publicaciones en general más breves que el promedio de los libros de texto o de consulta. Además se concentran en un tema y son autosuficientes. Con mucha frecuencia las monografías son publicaciones que tienden a formar parte de series con producción anual variable y sin una periodicidad establecida.
- d. *Ponencias de congresos*. Se trata aquí de la publicación de una colección de relatos de ponentes a reuniones de profesionales con un interés científico común que se congregan en un tiempo y lugar acordado. A menudo son de carácter regional, nacional o internacional.
- e. *Actas de Conferencias clínicas y anatómo-patológicas*. Son publicaciones de conferencias de médicos sobre sus observaciones de pacientes relacionando

su anamnesis y examen físico con los hallazgos de laboratorio, manifestaciones clínicas, diagnóstico diferencial y resultados de tratamientos. En el caso de las conferencias anatomo-patológicas se compara lo anteriormente descrito con los hallazgos de la autopsia, buscando explicaciones a posibles discrepancias entre ambas situaciones.

- f. *Informes Técnicos*. Estos son publicaciones de extensión habitualmente breve, sobre temas biomédicos u otros asuntos científicos relacionados. A veces los informes técnicos asumen el papel de «puestas al día» sobre un tema (enfermedad, condición humana o tecnología en uso), producto de la discusión de consenso de expertos.
- g. *Anuarios*. Consisten en la selección, compilación, ordenación y publicación de información general, estadística, técnica o científica de instituciones. Su frecuencia -como su nombre lo indica- es anual; a veces cumplen con una función estatutaria y de memoria técnica de la institución. El valor científico de lo que publican es superficial o ausente.
- h. *Boletines*. Suelen ser publicaciones periódicas de carácter informativo y técnico, de periodicidad variable. Su valor como vehículo de comunicación científica suele ser bajo, salvo aquellas excepciones en que se denomina «Boletín» a una verdadera publicación científica. Tal es el caso del *Bulletin of WHO* y del *Boletín de la Oficina Sanitaria Panamericana*.

2. Estructura de las revistas biomédicas

Una revisión atenta de las cinco revistas más frecuentemente consultadas en todo el mundo e incluidas en las bases de datos Index Medicus y Medline muestra que su contenido es variado con diversas secciones no siempre bien delimitadas.³

La estructura de estas prestigiosas publicaciones es como sigue:

The New England Journal of Medicine. (Semanal). Artículos originales, imágenes en medicina clínica, artículo especial, artículo de revisión, caso de los registros del Massachusetts General Hospital, editoriales, correspondencia, notas ocasionales, resección de libros, noticias, información para autores.

The Lancet. (Semanal). Editorial, comentario, artículos, informes breves, sesiones clínicas, artículo de revisión, punto de vista, libros, noticias, cartas al editor, información a los autores.

Annals of Internal Medicine. (Bimensual). Artículos, comunicación breve, revisiones, historia de la medicina, ser médico, editoriales, cartas, literatura de medicina, información a los autores.

British Medical Journal. (Semanal). Editoriales, noticias, artículos, práctica general, educación y debate, nota necrológica, cartas, sondeos, medicina y medios de comunicación, opinión personal, medicina y libros, minerva.

The Journal of the American Medical Association. (Semanal). Contribuciones originales, comunicaciones especiales, editorial, apéndices, resúmenes, cartas, libros, instrucciones para los autores.

En lo que respecta a las revistas biomédicas españolas, una selección de las mismas muestra los siguientes contenidos:

Medicina Clínica. (Semanal). Originales, editoriales, notas clínicas, artículo especial, conferencias de consenso, cartas al director, información para los autores, la imagen de la semana.

Gaceta Sanitaria. (Bimestral). Editorial, originales, artículo de revisión, cartas al director, información bibliográfica, actividades científicas, normas de publicación.

Atención Primaria. (Mensual). Editorial, originales, originales breves, formación continua, cartas al director.

Archivos Españoles de Urología. (Mensual). Urología experimental e investigación, urología general y urología oncológica, urología infantil, casos clínicos, comunicaciones breves y noticias.

Actas Urológicas. (Mensual). Artículos originales, revisiones conjuntas, investigación experimental y notas clínicas.

Revista de Calidad Asistencial. (Trimestral). Editorial, originales, comunicaciones, opinión, noticias, crítica de libros, normas de publicación.

Oncología. (Mensual). Revisiones, originales, notas clínicas.

Revista Clínica Española. (Mensual). Editoriales, originales, notas clínicas, revisiones, profesor invitado, medicina en imágenes, examínese de..., y cartas al director.

Revista de Sanidad e Higiene Pública. (Bimestral). Editorial, colaboraciones especiales, originales, actividades científicas.

El examen de la estructura de las revistas antes mencionadas muestra la presencia de algunas secciones comunes a la mayoría de ellas, las cuales se comentan someramente.

- a. *Artículos originales, contribuciones originales, originales, artículos.* Constituyen la forma predominante de publicación en las revistas biomédicas y de salud. Contienen información de tipo primario, es decir, primera divulgación de resultados originales de una investigación.
- b. *Artículos de revisión, revisiones, revisiones conjuntas.* En ellos se examinan áreas particulares de un trabajo o de un tema especial con el fin de informar sobre los avances que dicho tema ha tenido en un período de tiempo determinado. Las revisiones se deben llevar a cabo con la mayor seriedad científica por parte del investigador, quien revisa trabajos publicados y los selecciona con algún criterio explícito. El autor debe consultar directamente las fuentes, reuniendo elementos útiles, proponiéndolos, analizándolos, sintetizándolos y criticándolos.
- c. *Comunicaciones breves, informes breves, originales breves.* Como su nombre lo indica se trata de la publicación de datos de interés sobre un estudio limitado o de un informe de progreso de una investigación en desarrollo. A veces contienen los resultados preliminares de una investigación sobre un tema de impacto y actualidad.
- d. *Ensayos clínicos.* Informan sobre estudios planificados y usualmente controlados sobre la eficacia, la seguridad y los esquemas óptimos de dosis de una droga con aplicación, terapéutica, diagnóstica o profiláctica. También puede tratarse de una técnica, dispositivo o aparato lo cual se efectúa en seres humanos seleccionados de acuerdo a criterios predeterminados de elegibilidad y que se someten a observación para evidencia predefinida de efectos favorables o desfavorables.
- e. *Casos clínicos.* Esta es una forma muy popularizada de publicación en casi todas las revistas biomédicas. Obedecen al interés de los profesionales de compartir su experiencia de diagnóstico y tratamiento con pacientes individuales o pacientes agrupados alrededor de un mismo diagnóstico o problema clínico. El hecho de que los casos clínicos no sean una muestra representativa de la población los priva de validez externa en términos de poder extrapolar los resultados a un universo. Sin embargo los estudios de casos han hecho una contribución muy significativa al progreso de la medicina cuando se han realizado con el debido rigor metodológico.

- f. *Otras secciones.* Tal es el caso de la editorial, cartas al editor, informes de congresos y jornadas, noticias y reseñas de libros, las cuales cumplen con importantes objetivos editoriales pero que no contribuyen a la comunicación científica propiamente tal.

De todo este conjunto de elementos que constituyen el contenido de las revistas, el artículo científico original es sin duda el más importante y el de más difícil elaboración.

B. LAS ETAPAS DEL PROCESO DE INVESTIGACIÓN Y LA ESTRUCTURA DEL ARTÍCULO CIENTÍFICO ORIGINAL

M. A. Bobenrieth Astete

1. Concepto de investigación *

Investigación deriva etimológicamente de los términos latinos *in* (en, hacia) y *vestigium* (huella, pista). De ahí que su significación original es "hacia la pista" o "seguir la pista"; buscar o averiguar siguiendo algún rastro.

El ser humano es curioso por naturaleza; observa y tiende a indagar para comprender. Se interesa por la búsqueda de algo desconocido a partir de un vestigio conocido (huella, rastro o indicio), con el fin de describir, explicar o inferir. La investigación satisface -en parte- esta necesidad de conocer.

Conocer es averiguar, mediante el ejercicio de las facultades intelectuales, la naturaleza, las cualidades y las relaciones de las cosas o fenómenos. Conocimiento es entendimiento, inteligencia y razón natural. Tanto el científico como el detective, como el periodista se plantean y tratan de contestar a preguntas tales como: qué, quién, por qué, cómo, cuándo, dónde.

a. Definiciones de investigación

A continuación presentamos trece definiciones seleccionadas de otras tantas publicaciones con carácter de libros de texto, libros de consulta o manuales metodológicos.

- (1) Indagación sistemática, empírica, controlada y crítica de proposiciones hipotéticas teóricas acerca de relaciones supuestas que existen entre fenómenos (Kerlinger FN., 1979).

* Para información más detallada sobre este tema, sugerimos revisar la publicación "Investigación Científica en Salud. Desde el Problema hasta la Propuesta" (Documento Técnico N° 9. EASP, 1996) por Manuel Alejandro Bobenrieth Astete, editor.

- (2) Indagación, registro y análisis de evidencias con el propósito de ganar conocimiento (Ross R., 1983).
- (3) Búsqueda cuidadosa y diligente orientada al descubrimiento e interpretación de hechos, o la aplicación práctica de teorías nuevas (Webster D., 1985).
- (4) Estudio sistemático para descubrir hechos o teorías de análisis a fin de obtener respuestas válidas a preguntas planteadas o soluciones a problemas identificados (Notter LE, Hott JR., 1992).
- (5) Método sistematizado para obtener nuevo conocimiento, el cual puede verificarse y generalizarse más allá del caso o de la muestra estudiada (Brink PJ, Wood MJ., 1994).
- (6) Indagación sistemática que utiliza métodos científicos ordenados para responder preguntas y/o solucionar problemas (Polít DF, Hungler B., 1993).
- (7) Indagación sistemática que se acomete para solucionar problemas, responder preguntas y generar nuevo conocimiento. (Massey VH., 1995).
- (8) Serie de pasos que intentan dar una respuesta lógica a una pregunta específica (Baena G., 1993).
- (9) Manera por la cual resolvemos problemas complejos en nuestro intento de echar atrás las fronteras de la ignorancia humana (Leedy PD., 1992).
- (10) Averiguación sistemática y diligente para validar y refinar conocimiento existente y/o generar conocimiento nuevo (Burns N, Grove SK., 1995).
- (11) Estudio de un problema en pos de un objetivo determinado mediante el empleo de métodos precisos, con debida consideración al control adecuado de factores distintos de la variable bajo investigación y seguido de análisis, de acuerdo a procedimientos estadísticos aceptables (Comité de Investigación. Asociación Americana de Salud Pública., 1991).
- (12) Estudio científico y diligente a fin de establecer hechos y de analizar su significado. Inherente en tal estudio está un enfoque ordenado con registros exactos (Taber C., 1992).
- (13) Indagación diligente y examen de datos, informes y observaciones en búsqueda de hechos o principios (Mosby M., 1994).

b. Análisis de contenido

Una forma clara y atractiva de comprender el concepto de investigación es analizar el contenido de las definiciones; esto permitirá encontrar respuestas a cuatro preguntas claves: ¿Qué se hace (averiguación)? ¿Sobre qué se hace (objeto de estudio)? ¿Con qué propósito se hace (finalidad)? ¿Cómo se hace (procedimiento)?

- (1). Con respecto al qué se hace, la respuesta es indagar, buscar, estudiar, descubrir.
- (2). En lo relativo a sobre qué se hace, la respuesta incluye hechos y fenómenos; problemas empíricos, complejos, determinados y específicos.
- (3). En lo que respecta a con qué propósito se hace, la respuesta abarca solucionar problemas, aplicabilidad práctica, utilidad, explicar relaciones causa-efecto y generar nuevos conocimientos.
- (4). Referente a cómo se hace, la respuesta engloba métodos, sistemas, control, serie de pasos, orden, lógica, precisión, diligencia y ética.

2. Concepto de científico

a. Etimología

Científico proviene del latín "scientificus", adjetivo que significa perteneciente o relativo a la ciencia. Este término reconoce su origen en el vocablo latino "scientia" que se define como el conocimiento cierto de las cosas por sus principios y por sus causas. Hoy en día ciencia se homologa a un cuerpo de doctrina metódicamente formado y ordenado que constituye un ramo particular del saber humano. Aunque etimológicamente "ciencia" equivale a "el saber", hay saberes que no pertenecen a la ciencia; por ejemplo, el saber que se califica de común, ordinario, vulgar o cotidiano.

b. Tipo de conocimiento

Parece necesario, en consecuencia, precisar qué tipo de saber, qué tipo de conocimiento es el científico y distinguir entre el conocimiento científico y el conocimiento cotidiano. Surgen aquí algunas preguntas que han preocupado a la humanidad especialmente en la época moderna: a) ¿Qué es el conocimiento?; b) ¿Es posible el conocimiento?; c) ¿En qué forma debe expresarse el conocimiento?; d) ¿Cómo puede establecerse la verdad o falsedad del conocimiento?

Conocimiento es entendimiento, inteligencia, razón natural. Conocer es averiguar, por el ejercicio de las facultades intelectuales, la naturaleza, las cualidades y las relaciones de las cosas.

c. Teoría del conocimiento

Desde Emmanuel Kant, filósofo alemán del siglo XVIII, el problema del conocimiento comenzó a ser objeto de la teoría del conocimiento como disciplina filosófica especial, la cual centra su atención en cuatro aspectos fundamentales: a) La descripción del fenómeno del conocimiento o fenomenología del conocimiento; b) La cuestión de la posibilidad del conocimiento; c) La cuestión del fundamento del conocimiento; d) La cuestión de las formas posibles del conocimiento.

Escapa al objetivo eminentemente práctico de esta obra el abrir debate sobre la teoría del conocimiento como disciplina filosófica especial. Remitimos al lector interesado en la doctrina, los fundamentos y los métodos del conocimiento científico a informarse sobre este tema en algunas de las obras de epistemología disponibles en el mercado de libros de texto y de consulta.

3. Características del conocimiento científico

a. Fáctico

El conocimiento científico parte de hechos de la realidad, que se generan de alguna forma en la experiencia sensible. Para que un hecho pueda ser sujeto de investigación científica debe poder ser observable, registrable y medible. Marie Curie, física francesa de origen polaco, Premio Nobel de física en 1903 y de química en 1911, escribía: "En la ciencia hemos de interesarnos más por los hechos que por las personas".

b. Metódico

El conocimiento científico procede en forma organizada. Planea lo que persigue averiguar, la forma de obtenerlo, de analizarlo e interpretarlo. Dispone de procedimientos tales como la inducción, la deducción y la analogía. Pedro Laín Entralgo, médico, académico y escritor español contemporáneo declara: "El buen investigador ha de tener a la vez ideas nuevas, buena formación metodológica y buena técnica de trabajo. A los malos investigadores les falta siempre, por lo menos, uno de estos tres requisitos".

c. Analítico

El conocimiento científico distingue y separa las partes de un todo hasta llegar a conocer sus elementos, las relaciones entre éstos y los principios que rigen estas relaciones. Ivan Pavlov, fisiólogo y médico ruso, Premio Nobel de fisiología en 1904, declaraba a comienzos de este siglo: "Toda la ciencia positiva se ha edificado sobre el análisis y siempre, sin excepción, ha empezado por él".

d. Trascendental

El conocimiento científico parte de los acontecimientos pero va más allá de los hechos; se eleva por encima de eventos y situaciones aisladas hasta traspasar los límites de la experiencia posible. El conocimiento científico tiende a construir teorías, es decir, a generalizaciones abstractas las cuales presentan una explicación sistemática acerca de la relación entre fenómenos.

El conocimiento científico busca la generalización, la comprensión global de los fenómenos y de la forma en que éstos se interrelacionan. La capacidad para generalizar los resultados de la investigación es un criterio importante para estimar la calidad y el alcance de una investigación. Santiago Ramón y Cajal, médico e histólogo español, Premio Nobel de fisiología en 1906 nos recordaba: "Útil será recordar que todos los grandes investigadores han sido fecundos creadores de hipótesis; sin embargo cuán difícil es dictar reglas para imaginar hipótesis. Quien no posea cierta intuición del encadenamiento causal, cierto instinto adivinatorio para columbrar la idea en el hecho y la ley en el fenómeno, pocas veces dará, cualquiera que sea su talento de observador, con una explicación razonable".

e. Explicativo

A diferencia del conocimiento cotidiano, el conocimiento científico no se conforma únicamente con los hechos tal como se dan; investiga su origen, su asociación y, de ser posible, busca sus causas. Procura explicar los hechos en términos de teorías, de principios y de leyes. Bertrand Russell, filósofo y matemático inglés, Premio Nobel de literatura en 1950, explicaba: "Para llegar a establecer una ley científica existen tres etapas principales: la primera consiste en observar los hechos significativos; la segunda, en sentar hipótesis que si son verdaderas expliquen aquellos hechos; la tercera, en deducir de estas hipótesis consecuencias que puedan ser puestas a prueba por la observación".

f. Verificable

Los resultados y conclusiones del conocimiento científico deben someterse a prueba mediante observación y, de ser posible, experimentación. No puede aceptarse como cierto nada que escape a las pruebas de validez. Francis Bacon, filósofo y estadista inglés del siglo XVII, mencionaba: "Leer no para contradecir ni refutar, no para creer y aceptar como plenamente garantizado..., sino para calibrar y ponderar".

g. Simbólico

El conocimiento científico utiliza el lenguaje cotidiano pero va más allá; crea su propio lenguaje con palabras, signos y símbolos que suelen someterse a reglas.

h. Claro y preciso

Los conceptos científicos, los problemas objeto de estudio, las variables a investigar, las hipótesis, el diseño se definen sin vaguedad, sin ambigüedad. El lenguaje es directo y utiliza las palabras justas y las expresiones que corresponden exactamente a la idea que se quiere expresar. Azorín, escritor español de la Generación del 98, maestro de la naturalidad, la sobriedad y la brevedad de sus períodos oracionales, escribía: "En eso creo yo que está la perfección del estilo, en decir ni más ni menos lo que se quiere decir, y en decirlo con exactitud".

i. Predictivo

El conocimiento científico intenta explicar lo presente, lo pasado y aventura el futuro. La capacidad de predecir le sirve al científico para modificar acontecimientos mediante ciertas intervenciones, en beneficio de la sociedad (pronósticos médicos, predicción de epidemias, predicción de comportamientos, etc.). Auguste Comte, matemático y filósofo francés del siglo XIX, aseguraba: "Es necesario saber para prever, y prever para obrar".

j. Abierto (escéptico)

El conocimiento científico se basa en la noción de que cualquier proposición o afirmación, aunque provenga de fuentes importantes, es susceptible de cuestionamiento, de controversia. La ciencia trabaja con verdades provisionales: la verdad de hoy puede ser historia en el mañana. Jules Tannery, matemático y filósofo francés, en la alborada de este siglo afirmaba: "Es difícil encontrar un sabio que no sea escéptico; el sabio sabe tan poco, y sabe cuánto esfuerzo le cuesta ese poco". A su vez Bertrand Russell, nuevamente nos da su fundada opinión: "Ningún hombre de temperamento científico afirma que lo que ahora se cree en ciencia sea exactamente verdad; afirma que es una etapa en el camino hacia la verdad..."

k. Comunicable

El conocimiento científico conlleva una función informativa basada en datos sobre hechos observables, análisis e interpretación de los mismos, para beneficio del mayor número de personas.

l. Perseverante

El conocimiento científico es el producto no sólo de la inteligencia sino de la constancia del investigador, de su firmeza de ánimo, de su tenacidad para llevar las cosas adelante, de su perduración a través del tiempo. Marie Curie opinaba al respecto en forma enfática: "Hay que tener perseverancia y sobre todo confianza en sí mismo. Hay que creer que se está dotado para alguna cosa y que esa cosa hay que obtenerla

cueste lo que cueste". Por su parte Albert Einstein, físico alemán nacionalizado norteamericano, escribía: "Para nuestro trabajo son necesarias dos cosas: una de ellas es una persistencia infatigable; la otra la habilidad para desechar algo en lo que hemos invertido muchos sudores y muchas ideas".

m. Útil

El conocimiento científico es aprovechable, es beneficioso, es conveniente, es fructífero. Eventualmente puede ser lucrativo. Nuestro mundo actual, sin la ciencia y la tecnología inmersa en él, retornaría a la época primitiva.

4. Principios, procesos y métodos

a. Principios

La investigación científica en salud se funda en *principios*. Principio deriva del vocablo latino *arche* "el origen del que deriva el ser, la generación o el conocimiento". Parménides, filósofo griego del siglo IV-V a.C., señala el principio de identidad; Aristóteles, filósofo griego del siglo IV a.C. plantea los principios de contradicción, de causalidad y de finalidad; Wilhelm Leibniz, filósofo y sabio alemán del siglo XVII aporta el principio de razón suficiente. Para René Descartes, filósofo y sabio francés del siglo XVII- son primeras causas que llenan dos condiciones: son tan claros y evidentes que no cabe dudar de su verdad, por una parte, y son los fundamentos sobre los cuales depende el conocimiento de las demás cosas, por la otra. De entre estos principios destaca aquel de que el conocimiento es posible y aquel otro de que el fundamento de esa posibilidad es siempre la realidad; sea la realidad "sensible" (*empirismo*) o la realidad "inteligible" (*racionalismo*); sea que tome un punto de partida en el "objeto" (*realismo*) o en el "sujeto" (*idealismo*).

b. Positivismo

La investigación científica en salud se basa en el positivismo, doctrina y escuela fundada por Augusto Comte, matemático y filósofo francés del siglo XIX. Este sistema filosófico como teoría del saber se niega a admitir otra realidad que no sean los hechos y a investigar otra cosa que no sean las relaciones entre los hechos.

c. Método y metodología

Los términos método y metodología se usan a menudo indistintamente. De hecho bajo estos encabezamientos el investigador describe, con el mayor detalle que le permite el espacio, el diseño del estudio y los procedimientos actuales realizados en la

recogida, la tabulación, la presentación y el análisis de los datos. Aquí método/metodología se deben contrastar con la sustancia.

Creemos que sería muy útil el distinguir entre método como conjunto de técnicas y procedimientos para seleccionar sujetos, recoger, presentar y analizar datos, y metodología como el estudio de las suposiciones epistemológicas que guían el uso de un método. Métodos particulares (técnicas y procedimientos) pueden ser comunes a un número de enfoques de investigación las cuales se diferencian en forma importante una de otra en metodología.

d. Procesos sistemáticos

La investigación científica en salud se materializa en procesos sistemáticos los cuales se ordenan en etapas. Cualquier sistema de conocimientos se basa en el empleo de métodos apropiados; el sistema de conocimientos en clínica y en salud pública no hace excepción.

e. Proceso secuencial y circular

La investigación científica en salud abarca las formas de un proceso el cual posibilita el desentrañar la sustancia de la investigación. De acuerdo a las actuales tendencias del realismo y del pragmatismo, toda la realidad se da bajo el aspecto de un proceso. En el caso de las ciencias de la salud este proceso es -a la vez- lineal, secuencial y circular. Lineal por cuanto sigue una dirección: desde la pregunta (hipótesis) hacia la respuesta (conclusión); circular por cuanto cada pieza de información, cada evidencia factual origina una instancia de reflexión y de revisión de etapas anteriores.

Además la investigación es tanto un proceso lineal como circular por cuanto con base en la observación empírica (experiencia) establece, completa y reforma la teoría; a su vez, con base en la teoría intenta explicar una zona de la realidad.

5. Las etapas del proceso de investigación en salud

El proceso de la investigación en salud se compone de un conjunto de etapas o fases sucesivas de la serie de fenómenos a ocurrir en el espacio y en el transcurso del tiempo de la investigación.

Una revisión atenta de los diversos libros de texto y de consulta disponibles hay sobre metodología de investigación muestra -más allá de la diversidad de estructura y de nomenclatura- que el método científico se desarrolla en cinco grandes etapas consecutivas: 1) Etapa conceptual o etapa teórica; 2) Etapa empírica o etapa práctica; 3)

Etapa de planeación o etapa de diseño; 4) Etapa interpretativa o etapa del significado; 5) Etapa de diseminación o etapa de publicación.

A continuación se presentan -en forma sucinta- las cinco etapas mencionadas con los diversos pasos que las componen y la enumeración de las actividades que incluye cada paso (Ver esquema IV.I).

1. ETAPA CONCEPTUAL (TEÓRICA)

Paso 1 Definición del problema de investigación

(a) Desde el tema general al problema específico. (b) Concepto de problema. (c) Fuentes de problemas. (d) Factores en la selección del problema. (e) Formulación del problema de investigación.

Paso 2 Revisión bibliográfica

(a) Objetivos. (b) Momentos en que ha de llevarse a cabo. (c) Búsqueda y recuperación de la información. (d) Lectura crítica y evaluación de las fuentes.

Paso 3 Desarrollo del marco teórico/conceptual

(a) El problema dentro de un contexto teórico. (b) Los marcos conceptuales. (c) Los modelos.

Paso 4 Definición de variables

(a) Concepto. (b) Clasificación: Cualitativa/Cuantitativa; Discreta/Continua; Nominal/Ordinal; Independiente/Dependiente; Intervinientes/De control. (c) Medición: Frecuencia (incidencia/prevalencia); Asociación (riesgo relativo/riesgo atribuible); Impacto potencial. (d) Indicadores.

Paso 5 Formulación de hipótesis (objetivos)

(a) Concepto. (b) Criterios. (c) Clasificación: De trabajo/nula; Conceptual; Alternativa; Sencilla/compleja; Direccional/no direccional.

2. ETAPA DE PLANIFICACIÓN (DISEÑO)

Paso 6 Selección del diseño. Criterios

(a) Grado de intervención: Experimental; Cuasiexperimental; No experimental. (b) Tiempo o periodo: Transversal o seccional (prevalen-

cia); Longitudinal (incidencia). (c) Unidad de análisis: Individuos (casos); Muestras; Comunidades (ecológicos). (d) Nivel de análisis: Nivel I Descriptivo; Nivel II Analítico; Nivel III Inferencial.

Paso 7 Identificación de la población (universo)

(a) Población accesible. (b) Población blanco. (c) Medición: Frecuencia (incidencia/prevalencia); Asociación (riesgo relativo/riesgo atribuible); Impacto potencial. (d) Indicadores.

Paso 8 Determinación de la muestra

(a) Tamaño. (b) Representatividad: Muestreo probabilístico/no probabilístico; Muestras relacionadas/independientes; Error estándar.

Paso 9 Realización del estudio piloto o de prueba

(a) Objetivo. (b) Características. (c) Indicaciones.

3. ETAPA EMPÍRICA (PRÁCTICA)

Paso 10 Fuentes de información

(a) Documental (Primaria/secundaria). (b) De campo.

Paso 11 Recolección de datos

(a) Observación. (b) Experimentación. (c) Documentos. (d) Análisis de contenido. (e) Cuestionarios. (f) Entrevistas.

Paso 12 Procesamiento de información

(a) Edición. (b) Categorización. (c) Codificación. (d) Tabulación.

Paso 13 Presentación de información

(a) Texto. (b) Tablas. (c) Gráficos. (c) Otras ilustraciones.

Paso 14 Análisis de la información

(a) Niveles o escalas de medición: Nominal; Ordinal; Intervalar; De razón. (b) Análisis descriptivo: Cuantitativo; Cualitativo. (c) Análisis inferencial: De la estadística al parámetro; Ley de los grandes números; Errores tipo I (α) y tipo II (β); Estadística paramétrica y no paramétrica; Medidas de correlación y de regresión.

4. ETAPA INTERPRETATIVA (SIGNIFICADO)

Paso 15 Validez

Validez interna: (a) Variable dependiente (efecto), como consecuencia de la variable independiente (causa). (b) Papel de los sesgos, de variables extrañas (de confusión) y/o del azar. **Validez externa:** (a) Generalización de los resultados de la muestra al universo de donde provino la muestra. (b) Generalización de los resultados a otros universos. (c) Desarrollo de una teoría.

Paso 16 Importancia y novedad

(a) Aspectos relevantes. (b) Aspectos novedosos, originales.

Paso 17 Alcance

(a) Limitaciones: sesgos de selección, de información, de clasificación, de confusión. (b) Hallazgos novedosos y hallazgos confirmatorios; Hallazgos útiles.

Paso 18 Conclusión

(a) Respuesta (conclusión) a la pregunta de la investigación (hipótesis/objetivos) (b) Respuesta basada en los resultados. (c) Resultados basados en la metodología. (d) Metodología basada en el diseño. (e) Diseño basado en los objetivos/hipótesis.

5. ETAPA DE DISEMINACIÓN (PUBLICACIÓN)

Comunicación del estudio dentro de la institución

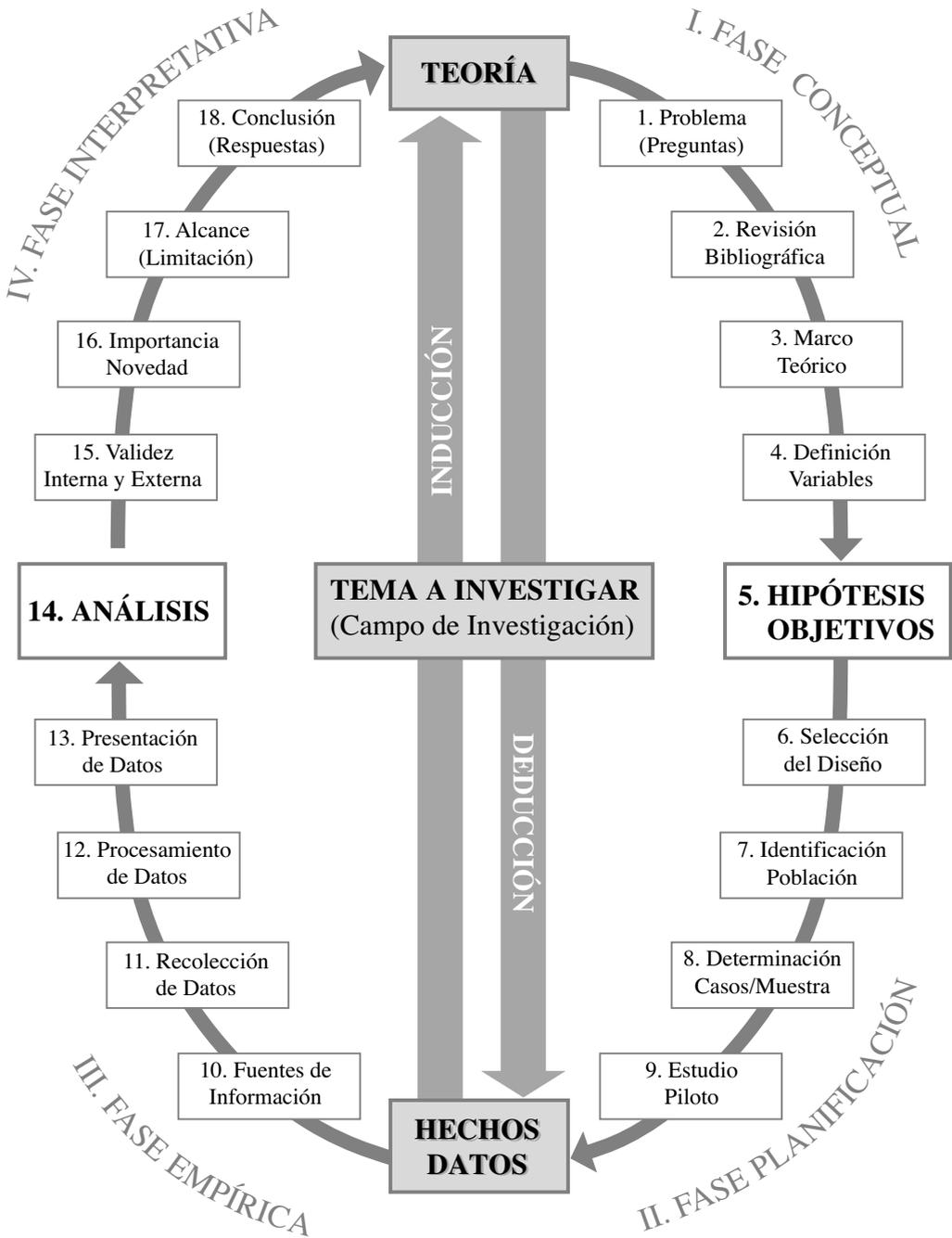
(a) Presentación oral. (b) Presentación de informe escrito.

Presentación del estudio a congresos y/o jornadas

(a) Presentación oral. (b) Presentación de ponencia/conferencia escrita.

Publicación de la investigación

Artículo científico original; Artículo especial; Comunicación breve; Caso clínico; Artículo de revisión.



Esquema IV.1. Modelo general del proceso de investigación

6. Las etapas de la investigación y los elementos del artículo original

El lector interesado en una investigación publicada en un artículo científico original sólo dispone de la información que se describe, se explica, se debate y se concluye en el documento impreso. En efecto, el lector no ha participado en el estudio ni le consta nada de lo que sucedió en él. En consecuencia debe actuar como "detective científico" para deducir lo que aconteció a lo largo del proceso de la investigación apoyándose exclusivamente en los datos, en las evidencias tanto directas como circunstanciales que narra el autor del artículo.

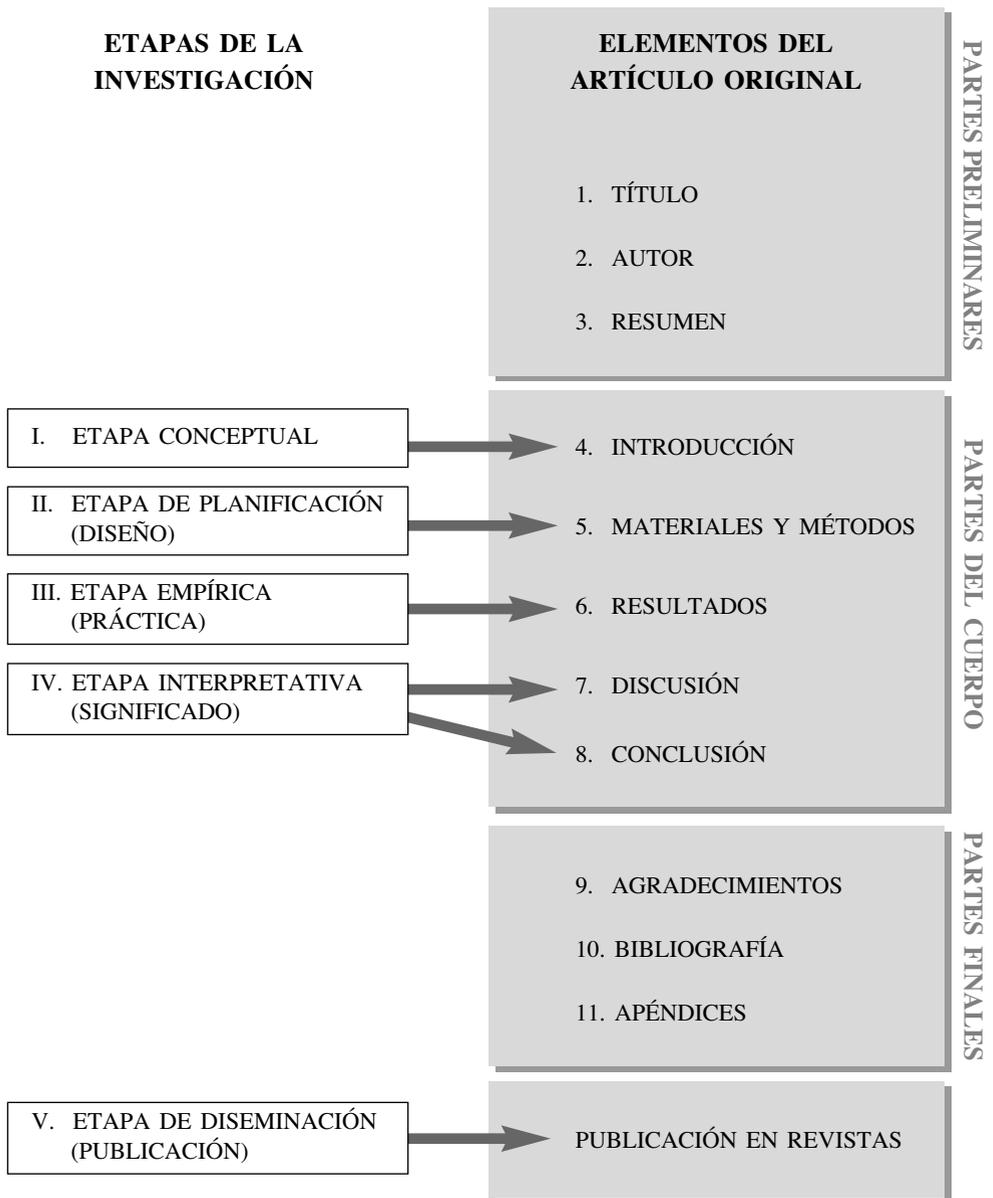
Las consideraciones antedichas plantean la necesidad imprescindible de que el artículo científico original se estructure siguiendo la secuencia lógica/psicológica de las etapas del proceso de la investigación. La comunidad científica internacional ha adoptado el esquema "IMRYD": Introducción, Material y métodos, Resultados Y Discusión/conclusión.

Cada uno de estos cinco elementos o partes del cuerpo del artículo científico original se corresponde a una de las etapas del proceso de la investigación. De esta manera:

- a. La Introducción contiene lo acontecido en la etapa conceptual o etapa teórica;
- b. Los Materiales y Métodos incluyen lo sucedido en la etapa de planificación o diseño;
- c. Los Resultados muestran los hallazgos de la etapa empírica o práctica;
- d. La Discusión y la Conclusión reflejan fielmente los análisis, las reflexiones, los debates, las polémicas y las deducciones de la etapa interpretativa o del significado.

El Esquema IV.2 "Las etapas de la investigación frente a la estructura del artículo original" muestra las relaciones descritas.

La Parte C de esta obra: "Partes del artículo científico original" describe y explica con detalle los once elementos o partes del artículo científico original.



Esquema IV.2. Las etapas de la investigación frente a la estructura del artículo original

C. PARTES DEL ARTICULO CIENTIFICO ORIGINAL

M. A. Bobenrieth Astete

Un artículo científico original es un escrito que contiene una descripción clara, concisa y completa de una investigación. Su finalidad es comunicar con fidelidad los métodos, los hallazgos y su interpretación, no como parte de un libro, sino como un todo acabado estructurado e interesante.⁴

Conviene anticipar que sus partes se pueden agrupar en tres: a) Partes preliminares: título, autor(es) y resumen; b) Partes del cuerpo: introducción, materiales y métodos, resultados, discusión y conclusión; c) Partes finales: agradecimientos, referencias y apéndice.

La clase de artículo científico original varía según el tema de la investigación y el diseño seleccionado para lograr los objetivos definidos. Así se los puede clasificar en artículos experimentales y no experimentales (también llamados «observacionales»); artículos de tipo descriptivo y de tipo analítico; artículos teóricos y artículos aplicados.

En relación con los artículos teóricos el término «teoría», en el sentido de conocimiento puro, se opone al término «práctica» la mayoría de las veces, entendida como cualquier clase de actividad orientada hacia el exterior (fuera del conocimiento mismo).⁵ De ahí que se hable de una exposición teórica, cuando tiende a plantear y relacionar determinados conceptos; práctica, -a su vez- cuando tiende a llevar a la acción dichos conocimientos. Los artículos teóricos pueden entenderse en sentido amplio y en sentido estricto; en sentido amplio, un artículo teórico es aquel que expone conocimientos organizados, considerándolos con independencia de toda aplicación aunque no opuesto a ella. En sentido estricto, un artículo teórico es aquel en el cual se expone un sistema de hipótesis entre las que se destacan las leyes acerca de un sector de la realidad.

De entre los artículos de tipo analítico cabe destacar los estudios correlacionales en los cuales se trata de determinar interrelaciones entre variables clínicamente significativas.

Las categorías de artículos mencionados no son necesariamente excluyentes ya que en una misma investigación se pueden mezclar algunas categorías. Los trabajos de investigación publicados en forma de artículos originales, han de elaborarse y presentarse siguiendo los principios y las normas descritas en los capítulos siguientes.

1. Título

Concepto y objetivos

Un título es inevitable en la primera página del artículo original. Lamentablemente con demasiada frecuencia este primer dato del manuscrito que obtiene el editor de la revista o el lector del artículo ya publicado, suele ser ampuloso, extenso, confuso, ambiguo, partido, con abreviaturas y siglas y en tono efectista.⁶ No es de extrañar con estos síntomas y signos, que la «supervivencia» del artículo es mínima.

Un buen título debe tener atractivo y garra; el título tiene que captar la mirada y la atención del lector que hojea la revista; un buen título debe despertar curiosidad y estimular el interés por leer el artículo. El profesional de la salud, aun aquellos que están suscritos a una revista, no lee todos los artículos. La decisión de qué artículo leer es selectiva y para tomar esta decisión el lector se basa comúnmente en el título del artículo. Luego si éste le resulta atractivo se fijará en los autores para aterrizar en el resumen.

Si el artículo original, al igual que cualquier otro documento científico, necesita de un título, éste debe ser indicativo del tema estudiado identificando con precisión el tema principal del artículo. Obviamente no debe abarcar información no incluida en el texto del artículo. El título debe valorarse por la información que contiene y por su claridad y exactitud; en consecuencia debe contener los elementos principales que el texto del artículo intenta comunicar. Si el título no indica que el contenido encaja en el margen de interés del lector, cuando en efecto sí lo hace, el autor habrá perdido un lector y el lector habrá perdido un artículo interesante. A su vez si el título indica que el contenido encaja en el campo de interés del lector y resulta no ser cierto, entonces el lector habrá perdido su tiempo, se frustrará y se molestará.⁷

No es objetivo del título el destacar cada uno de los aspectos interesantes de un artículo. En consecuencia los autores deben ejercitar su mejor juicio para incluir palabras claves correspondientes a los 2 o 3 aspectos más relevantes del estudio. El título es una «etiqueta» clara y útil que anuncia el trabajo mismo de investigación. El título no debe sonar como una afirmación tajante y dogmática. No se acepta hoy en día indicar conclusiones en los títulos; para eso está el resumen del artículo y la sec-

ción de discusión del mismo. La utilización de palabras claves (o descriptores) subrayan el contenido significativo del estudio para que resulte comprensible, localizable y recuperable. La importancia del título se acrecienta por el hecho que es un elemento de lo que se conoce como la «parte bibliográfica» de un artículo. Los otros son la lista de autores, nombre de la revista, año de publicación, número de volumen y páginas que abarcará.

El título proporciona información que se utiliza para los sistemas informatizados de almacenamiento de datos. Los sistemas computarizados de búsqueda de literatura biomédica y de salud, tales como MEDLARS, Chemical Abstracts Services, CA Bibliofile, etc., se han desarrollado con la suposición de que los artículos relevantes puedan localizarse sobre la base de palabras claves en sus títulos.

Numerosos servicios bibliográficos dependen fuertemente de los títulos para clasificar, seleccionar e indexar los artículos publicados. Las referencias bibliográficas en libros y revistas incluyen -entre otros elementos- los títulos de las publicaciones, a fin de facilitar la verificación de la información mencionada y/o la eventual profundización en el tema.

Un número creciente de sistemas importantes de alertas bibliográficas, tales como *Current Contents*, *Science Citation Index*, *Chemical Titles*, etc., al igual que servicios de distribución selectiva de resúmenes, dependen para su efectividad de la adecuación y exactitud de los títulos de los artículos recientemente publicados.

El «*Permuterm Subject Index*» del *Science Citation Index* es un ejemplo importante de una fuente de información en la cual los términos de los títulos incorporados juegan un papel crucial. Este índice se construye usando las permutaciones posibles de pares de palabras significativas de títulos de reciente publicación, listadas en orden alfabético. La premisa implícita de este sistema es que muchos científicos escrutarán las listas en busca de palabras centrales a su interés de investigación, lo cual les permitirá ubicar publicaciones nuevas relevantes a su trabajo. Pares de palabras claves suelen ser más definitivas a este respecto que palabras claves solas.

Estructura

Con frecuencia los títulos son de difícil comprensión debido al uso de palabras incorrectas y/o a un orden equivocado de las palabras. La redacción de un buen título requiere no solamente de la utilización de *palabras correctas* sino también de *sintaxis correcta*.⁸ La sintaxis trata de la forma de enlace o asociación entre palabras, conforme a las reglas de la gramática española.

Desde un punto de vista de la estructura, los títulos se refieren a investigaciones de tipo causal o a investigaciones de tipo descriptivo. En investigación de tipo causal los títulos adoptan, en general, una de las dos formas siguientes:

- Y como efecto (función) de X
- X como causa de Y

En investigación de tipo causal, Y se refiere a la variable dependiente objeto de medición; X se refiere a la variable independiente que manipula el investigador.

En investigación de tipo descriptivo, los títulos informan al lector acerca de:

- sujetos (grupos) que se estudian.
- características de los sujetos bajo estudio.

No es aceptable la combinación de un título con subtítulo (Título. Subtítulo). Tampoco es aceptable un título partido (Título: subtítulo). Los subtítulos tienden a omitirse porque suenan pedantes, requieren puntuación y dificultan la indización. De hecho los términos claves del subtítulo siempre pueden incorporarse en el título, muchas veces al comienzo del mismo.

No deben emplearse abreviaturas, jerga o terminología insólita, anticuada. Tampoco deben emplearse vocablos o locuciones que solo tienen uso en un área restringida. Tampoco es aceptable el uso de fórmulas químicas, nombres de fantasía de drogas y otros productos patentados, en lugar de nombres genéricos.⁹

A veces es conveniente sugerir un «titulillo», el cual no es sino una versión abreviada del título. Éste suele usarse ocasionalmente como «título de página» de la revista correspondiente. Habitualmente el titulillo va en la parte superior de las páginas del lado izquierdo; la parte superior de las páginas del lado derecho se reservan para el «título de la revista».

Estilo

Desde un punto de vista del estilo, el título tiene que ser claro, atractivo, directo y exacto. Debe utilizar términos precisos, específicos, unívocos y tiene que ser conciso y suficientemente descriptivo. Los títulos oscuros o equívocos confunden al lector y lo desalientan a proseguir a leer el resumen.

Se debe incluir el menor número posible de palabras para describir adecuadamente el contenido del artículo. La experiencia aconseja que un máximo de 15 palabras son suficientes para la inmensa mayoría de los títulos. Los títulos demasiado

largos no son más significativos que los cortos ya que de hecho suelen contener palabras superfluas e inútiles

A su vez los títulos demasiado cortos de estilo telegráfico tampoco son aceptables por ser incompletos, demasiado vagos, ambiguos e inespecíficos. Se trata de ser breve y conciso sin ser críptico, enigmático o incompleto.

El título debe redactarse como una frase en tono afirmativo. No es aceptable una frase en tono interrogativo.

Faltas frecuentes

DE CLARIDAD

Los conceptos y razonamientos son de difícil comprensión por:

- a. Uso de términos ambiguos, vagos.
- b. Uso de jerga o jerigonza.
- c. Uso de abreviaturas o siglas.
- d. Uso de palabras incorrectas.
- e. Utilización de sintaxis incorrecta.
- f. Uso de títulos partidos.
- g. Uso de títulos con subtítulos.
- h. Uso de títulos y/o subtítulos efectistas. Ejemplos:
 - «Cirugía urológica en el octogenario: ¿Demasiados prejuicios?»
 - «Aspectos clínicos de la hiperplasia benigna de la prostata: los tiempos cambian.»

DE CONCISIÓN

Los conceptos no se expresan en forma acertada con el menor número posible de palabras.

- a. Títulos demasiado extensos.
- b. Títulos demasiado breves.
- c. Títulos con exceso de preposiciones y artículos.
- d. Uso de subtítulos innecesarios.

DE SOBREENPLICACION

La declaración o exposición repetitiva de un concepto que se da por supuesto los 25 ejemplos que se mencionan a continuación provienen de títulos de artículos publicados.

«Estudio sobre»....
«Estudio preliminar de»
«Nuevo estudio sobre»....
«Investigación acerca de»....
«Informe de»....
«Situación de»....
«Estado de la situación de»....
«Observación acerca de»....
«Observaciones preliminares sobre»....
«Conceptualización de»....
«Nuestra experiencia de»....
«Resultados de un estudio sobre»....
«Resultados de la investigación de»....
«Análisis de los resultados de»....
«Determinación de»....
«Análisis de la situación de»....
«Comportamiento de»....
«Controversia ante»....
«Notas sobre»....
«Un tratamiento nuevo para»....
«Valoración del»....
«Contribución al»....
«Consideraciones acerca de»...
«A propósito de»...
«Una aproximación a»...

Las palabras entre comillas resultan superfluas e inútiles; están de más por ser innecesarias y deben omitirse.

2. Autores

Concepto y objetivos

Las observaciones y conclusiones de un investigador son de su propiedad intelectual hasta que las presenta a la comunidad científica en forma de un trabajo publicado, habitualmente como un artículo original.

Es fundamental el establecer «criterios de autoría» de un artículo original. ¿Quién tiene derecho a aparecer como autor de un artículo científico? Para que una persona aparezca como autor en propiedad, debe reunir los siguientes requisitos: a) Haber hecho una contribución intelectual sustancial a la investigación; b) estar en condi-

ciones de presentar el trabajo y defender las críticas en reuniones de profesionales especialistas. En otras palabras debe poder asumir responsabilidad pública del contenido total del artículo, lo cual se expresa en poder explicar el porqué y el cómo se realizaron las observaciones; cuál fue el marco teórico y cómo se decidió el tipo de diseño de la investigación; cuáles fueron los hallazgos principales y cuál es su análisis y su significado.

De lo anterior se desprende que no califican como autores aquellas personas con una conexión escasa y circunstancial con el contenido y el proceso de la investigación. Se incluyen aquí autoridades, directores, jefes de servicios o departamentos que han autorizado y patrocinado la investigación, personas que han proporcionado ayuda económica, gente que facilitó locales y equipos para la investigación, personas que proporcionaron alguna ayuda técnica, etc.. A este último grupo de personas se les puede mencionar en la sección de «Agradecimientos» del artículo. El haber participado únicamente en la recolección de datos (encuestadores, etc.) tampoco da derecho a aparecer como autor.

Otra cuestión no siempre resuelta de forma satisfactoria es el orden de aparición de los nombres de los autores. Aun cuando no existen reglas mandatorias, sí existe -en general- un consenso en la comunidad científica internacional: el primer nombre corresponde al autor principal. Luego vienen los demás autores en orden decreciente de importancia de la contribución.

Estructura

Los últimos 40 años han sido testigos de una tendencia creciente a la *autoría múltiple* de los artículos originales publicados en revistas biomédicas. Esto se explica por el crecimiento de las ciencias y de sus publicaciones, por la complejidad progresiva de los estudios en el campo de la salud y de la enfermedad, por el aumento de la investigación en equipos multidisciplinarios, y por el «credencialismo» vigente.

Más allá de las razones descritas de la autoría múltiple, hay otra razón subyacente: el conceder crédito de autor a personas que no lo merecen. La situación antes descrita constituye la *autoría injustificada*, la cual es una infracción grave a la ética de publicar. A su vez la situación de negarle el crédito como autor a una persona que lo merece por ser responsable de una parte decisiva del contenido intelectual del artículo constituye una injusticia y se denomina *autoría incompleta*. La combinación de la autoría injustificada y la autoría incompleta constituye la *autoría irresponsable*, la cual es parte de los abusos de carácter ético que se producen en el campo de la publicación, junto con el error, el sesgo, el plagio y el fraude.¹⁰

Estilo

Las normas sobre autoría emitidas por el Comité Internacional de Editores de Revistas Médicas (CIERM) han logrado gran aceptación entre muchas revistas biomédicas de todo el mundo, incluidas algunas de las más prestigiosas. Por la trascendencia y utilidad de dichas pautas, vale la pena reproducirlas:

«AUTORIA.

Todas las personas designadas como autor deben cumplir ciertos requisitos para tener derecho a la autoría. Cada autor debe haber participado en el trabajo en grado suficiente para asumir responsabilidad pública por su contenido.

Para concederle a alguien el crédito de autor hay que basarse únicamente en su contribución esencial: a) La concepción y el diseño del estudio, o el análisis y la interpretación de los datos; b) La redacción del artículo o la revisión crítica de una parte importante de su contenido intelectual; c) La aprobación final de la versión que será publicada. Los requisitos a, b y c deben cumplirse siempre.

La participación que se limita a conseguir financiamiento o recoger datos no justifica que se le conceda a nadie el crédito de autor. Tampoco basta con haber ejercido la supervisión general del grupo de investigación. Toda la parte del artículo que sea decisiva con respecto a las conclusiones principales debe ser responsabilidad de por lo menos uno de los autores.

En un artículo de autor corporativo (colectivo) se debe especificar quiénes son las personas principales que responden del documento; a las demás personas que colaboraron en el trabajo se les debe conceder un reconocimiento por separado (véase Agradecimientos).

Los directores de revistas podrán solicitar a los autores que justifiquen la asignación de la autoría.»¹¹

Faltas frecuentes

- a. Incluir como autores a personas que no han hecho una contribución intelectual sustancial a la investigación y que no pueden defender públicamente el contenido del artículo en su totalidad (autoría injustificada).
- b. Incluir un número desproporcionadamente alto de autores (autoría múltiple).
- c. Excluir verdaderos autores (autoría incompleta)
- d. No respetar el criterio de magnitud de la contribución al decidir el orden de aparición de los autores.

- e. Usar iniciales en lugar de nombres completos de los autores
- f. Incluir grados académicos o posiciones jerárquicas de la burocracia institucional.
- g. No incluir la dirección postal completa del autor principal.

3. Resumen

Concepto y objetivos

El término «resumen» se usa en escritura científica como una expresión o declaración abreviada reducida a términos sucintos, precisos y esenciales de las ideas y conceptos más importantes del informe de una investigación.¹² Su confección es el reto más serio a la capacidad de síntesis del autor.

El crecimiento exponencial de las revistas biomédicas durante los últimos 50 años y del número de artículos originales publicados hizo necesario el incluir resúmenes que proporcionan información rápida a los profesionales ocupados. Esta parte debe describir las ideas más importantes y los contenidos sobresalientes del documento.

La mayoría de los profesionales leen solamente el título del artículo. Cuando el contenido de éste parece encajar en su campo de interés, algunos avanzan un paso más y leen el resumen. Si su lectura les causa una impresión positiva, entonces decidirán leer el artículo. Un buen resumen destaca lo novedoso. No es un mini artículo sino una selección de algunos aspectos importantes del estudio, acordados por consenso por la comunidad científica internacional; bien redactado produce una impresión positiva de inicio. Sin embargo redactarlo es tarea muy difícil, lo cual explica por qué la inmensa mayoría de los artículos publicados los traen incompletos y deficientes.

Un buen resumen debe ser inteligible, completo, exacto, objetivo y conciso; debe dar datos concretos y no compilar generalizaciones. Al igual que en el caso del título, debe evitar abreviaciones y fórmulas. Tampoco debe incluir referencias citadas en el texto; igualmente debe excluir tablas, gráficos u otras ilustraciones.

El resumen de un artículo publicado se diferencia del resumen de una ponencia a un congreso profesional, habitualmente enviado con anticipación a la fecha del evento. El resumen de una ponencia se rige por los mismos principios del resumen del artículo pero es habitualmente un poco más extenso y a veces llega a constituir una comunicación breve.

Los objetivos se pueden enunciar:

- a. Permitir que el lector identifique el contenido básico del artículo de forma rápida y exacta;

- b. Permitir que el lector determine la pertinencia (que viene a su propósito) y la relevancia (que es importante y significativo) de este contenido a su interés particular;
- c. Ayudar al lector a decidir si le interesa leer el documentos en su totalidad.

El resumen es una entidad independiente y, por lo tanto, debe ser autoexplicativo, autónomo. Debe ser inteligible por sí mismo, sin referencia al artículo, pero sin ser un sustituto de éste. La incorporación de los resúmenes en fuentes bibliográficas y bases de datos especializadas agrega importancia al resumen. De hecho el resumen, junto con el título y los autores es la parte del artículo que se incluye en los sistemas de información bibliográfica (Index Medicus, Excerpta Médica, etc..).

Estructura

La extensión del resumen es de 150 a 300 palabras (alrededor de un 5% de la extensión del texto del artículo). Esto obliga al autor a sopesar cuidadosamente cada frase y cada palabra que se incorpore a él.¹³

Hay dos tipos de resumen: el resumen descriptivo y el resumen informativo (también llamado a veces indicativo). a) El resumen descriptivo tiene por objeto presentar meramente el tema del artículo sin entrar a los aspectos sustantivos del mismo. Se utilizan con frecuencia en comunicaciones a conferencias, informes oficiales y artículos de revisión. No sirven para los propósitos del resumen de un artículo original, el cual contiene información de tipo primario. b) El resumen informativo tiene por objeto el condensar aquellos aspectos más importantes del estudio. El resumen informativo suele adoptar dos modalidades: el resumen tradicional y el resumen estructurado.

En su «forma tradicional» el contenido del resumen debe expresar de forma clara, concisa y secuencial los siguientes 4 elementos:

- a. Los objetivos (ó hipótesis) principales del estudio;
- b. Los métodos empleados, incluyendo el tiempo, el lugar, los sujetos de la investigación y la intervención efectuada;
- c. Los resultados (hallazgos) principales con valores numéricos precisos.
- d. Las conclusiones principales con sus límites de validez y los alcances del estudio.

Las inmensas dificultades que enfrentan las revistas biomédicas para conseguir resúmenes que incluyan espontáneamente estos 4 elementos fundamentales, por una

parte, y la necesidad de disponer de información exacta sobre el tipo de diseño de investigación con énfasis en aspectos metodológicos, por la otra, ha dado origen en revistas de medicina clínica a una nueva modalidad de resumen de artículo original denominado «resumen estructurado».

En su «forma estructurada» el resumen debe expresar en forma clara, concisa y secuencial los siguientes 8 elementos:

- a. Objetivo
- b. Diseño
- c. Contexto
- d. Sujetos (participantes, pacientes, etc..)
- e. Intervenciones
- f. Medición del desenlace
- g. Resultados principales
- h. Conclusión

Una comparación rápida del resumen tradicional frente al resumen estructurado muestra que son idénticos en los ítems primero, penúltimo y último (objetivos, resultados y conclusión). La verdadera diferencia está en el ítem b) «métodos empleados» del resumen tradicional, el cual se expande a cinco ítems en el resumen estructurado: diseño, contexto, sujetos, intervenciones y medición del desenlace.

Estilo

La redacción del resumen debe tomar en cuenta dos partes fundamentales: a) el fondo, el cual se asegura en la lógica del pensamiento y en el orden de las ideas; b) la forma, la cual se constituye en la manera de expresar gramaticalmente los pensamientos.

El resumen es una condensación de los conceptos principales del contenido del documento. Condensación significa reducir a menor extensión un escrito o discurso sin quitarle nada de lo esencial.

El término «resumen» no debe confundirse con otros términos relacionados, tales como: anotación, extracto y sumario. Una «anotación» es un comentario breve, una explicación sobre un documento, o sobre alguna parte de él. Un «extracto» es una o más partes de un documento, seleccionadas para representarlo. Un «sumario» es una breve condensación parcial del documento, habitualmente de una parte, capítulo o sección de él.

El estilo del resumen debe, en consecuencia, asegurar el cumplimiento de una serie de requisitos: una forma de expresión clara, concisa, objetiva e impersonal; el empleo del pasado verbal (tiempo pretérito), excepto en frases concluyentes, ya que relata un estudio realizado; una extensión máxima de 300 palabras; su ubicación en página separada, entre el título y la introducción; la exclusión de abreviaturas no explicadas; la exclusión de referencias bibliográficas; la exclusión de tablas y gráficos; la exclusión de información o conclusiones que no estén en el texto.

Faltas frecuentes

- a. Omiten elementos importantes.
- b. Destacan asuntos insignificantes.
- c. No es inteligible por sí mismo.
- d. No incluye los objetivos.
- e. No incluye los materiales y los métodos.
- f. No incluye los resultados más relevantes.
- g. No incluye las conclusiones.
- h. Incluye conclusiones no mencionadas en el texto del artículo.
- i. Incluye otra información que no figura en el texto del artículo.
- j. Utiliza abreviaturas y símbolos poco conocidas.
- k. Carece de claridad.
- l. Carece de precisión.
- m. Carece de concisión.
- n. Carece de sencillez.
- o. Carece de secuencia lógica; las frases y los párrafos no están ordenados adecuadamente.
- p. No es autoexplicativo, autónomo; no se entiende separado del artículo original.
- q. Utiliza nombres coloquiales o marcas registradas de productos químicos o farmacéuticos en lugar de denominaciones genéricas.
- r. Menciona localidades geográficas poco conocidas sin indicar denominación estatal ni país.

4. Introducción

Concepto y objetivos

El cuerpo del artículo científico original se inicia con la introducción. El lector espera encontrar allí descritos los diversos elementos de la fase conceptual de la investigación. La introducción responde al «qué» y al «por qué» de la investigación. Como el sentido etimológico lo indica, la introducción es la entrada del lector al tema. La introducción da cuenta de todo aquello que debe presuponerse objetivamente para la recta inteligencia de la cuestión, expuesta en el artículo.¹⁴

La introducción sirve de transición del mundo del lector al mundo del autor del artículo y prepara al lector para que lea el artículo con atención y con reflexión. Para ello el autor debe presentar prontamente el campo de estudio para pasar luego al problema de investigación. La información proporcionada deberá ser suficiente como para que aun el lector que no conozca el tema en detalle pueda ubicarse fácilmente en el trabajo. Una buena introducción despierta el interés del lector en conocer el resto del artículo.⁶

La exposición del *problema* incluye: a) Identificar el problema por observación, por estudio, o ambos; b) Describir el origen del problema. Se deben exponer los antecedentes del problema, valiéndose de la primera revisión bibliográfica; de esta manera el lector puede conectar los antecedentes conocidos con los hechos que se propone investigar; c) Destacar la magnitud y la importancia del problema; d) Diferenciar los diversos elementos del problema; e) Distinguir el problema general del problema de investigación.

Cualquier situación conflictiva no es objeto de investigación. El problema de investigación debe reunir 4 requisitos: a) Ser susceptible de observación y de medición; b) Requerir una respuesta o solución; c) Contribuir a la generalización de nuevos conocimientos y/o tecnologías; d) Tener una utilización práctica.

En la primera caracterización del problema se trata de averiguar qué se conoce del mismo, qué se ha investigado sobre el mismo y qué aspectos importantes permanecen aún desconocidos. Los criterios utilizados generalmente en la formulación del problema de investigación son: a) Expresar algún tipo de relación o asociación entre variables; b) Posibilitar la prueba empírica de las variables; c) Expresarse en una doble dimensión temporal y espacial; d) Formularse en forma de pregunta.

La averiguación sobre los *antecedentes del problema* conduce a la primera búsqueda bibliográfica. Ésta consiste en una revisión, selección y exposición breve de trabajos publicados, cuyos resultados corresponden al tema específico. La revisión bibliográfica reconoce la contribución de otros investigadores y muestra la continuidad lógica entre sus trabajos y la investigación objeto del artículo.

Para la discusión de la literatura se tienen en cuenta las siguientes indicaciones: a) Seleccionar citas directamente pertinentes al tema; b) Seleccionar citas actualizadas al tema (idealmente los últimos 10 años); c) Centrarse en publicaciones de tipo primario; d) Poner énfasis a los hallazgos, procesos metodológicos y conclusiones más significativas, sin incluir detalles secundarios; e) Evitar una revisión bibliográfica exhaustiva (Ver Apéndice II "El proceso de la revisión bibliográfica. Resumen esquemático").

La formulación del *marco teórico* permite: a) Delimitar lo que incluye la investigación; b) Visualizar las interrelaciones de los diversos elementos que componen el problema; c) Establecer las teorías que son pertinentes al problema, resumiendo los argumentos y datos más relevantes.

La *definición de términos* se realiza explicando aquellos términos claves para su fácil comprensión. Todos los términos que se aparten del uso convencional requieren una definición particular. La definición ha de ser clara y convenir solamente a lo que define. Las definiciones son indispensables; en caso contrario, el lector no podrá seguir con claridad el hilo del asunto. En otras palabras la falta de definiciones es el comienzo de la confusión; por esto es importante incluirlas en la introducción.

La *definición de variables* se utiliza preferentemente en los trabajos experimentales y cuasi-experimentales. Se deben indicar la variable independiente y la variable dependiente. Además las variables de control, las variables intervinientes y las variables de confusión.

El *objetivo* es el resultado concreto y previsto que se intenta obtener mediante la investigación. El objetivo definido no debe perderse en divagaciones sin relación directa con el problema ya que el trabajo debe indicar qué es exactamente lo que se pretende. En ningún momento el investigador debe separarse del objetivo propuesto; si surgen otras cuestiones afines se anotarán a medida que se desarrolla el estudio para su ulterior discusión. El objetivo de la investigación, en consecuencia: a) Enuncia un resultado unívoco, claro, preciso, factible y medible que se obtendrá una vez terminado el proceso de investigación; b) Define un estado o situación cuantificable en un lugar y tiempo determinado, que se intenta alcanzar como resultado de la investigación; c) Corresponde a la pregunta de investigación cuya respuesta constituirá la conclusión del estudio.

La *formulación de hipótesis* consiste en una afirmación de relaciones previstas entre dos o más variables bajo investigación. Una hipótesis es una explicación anticipada, una solución provisional del problema planteado. Esta solución se fundamenta en observaciones del propio autor y en los datos derivados de otras investigaciones publicadas. Las hipótesis conducen a estudios empíricos que buscarán confirmarlas o rechazarlas. La hipótesis recibe la prueba empírica a través de la recolección y el análisis de los datos.

Estructura

La estructura de la introducción debe corresponder a los conceptos y objetivos de este capítulo del artículo científico original. En consecuencia debe organizarse reflejando fielmente los elementos allí definidos.

El texto de la estructura debe comprender los siguientes elementos, en el orden que se indica:

- a. Enunciación del problema general.
- b. Revisión de los antecedentes del problema y examen de la literatura pertinente (fundamentos del estudio).
- c. Definición del problema de investigación.
- d. Establecimiento de un marco teórico.
- e. Definición de las variables.
- f. Definición operacional de términos claves.
- g. Definición de objetivos (o hipótesis) del estudio.
- h. Importancia y alcance del estudio.
- i. Limitaciones del estudio.

El problema general que llevó a la investigación surge de la observación espontánea o sistemática de los hechos. El problema de investigación es habitualmente un subproblema del problema general, es un aspecto o componente del problema general. En el planteamiento del problema debe indicarse su origen, cómo se entendía antes, cómo se entiende hoy, cuáles son sus elementos, cómo se relaciona con cuestiones afines y cómo se distingue de ellas. La enunciación del problema debe ser breve y clara. Puede plantearse en forma de pregunta, de tal manera que lo entienda el mayor número de lectores y no únicamente los especialistas.

En la revisión de los antecedentes del problema, se exponen brevemente los trabajos más importantes relacionados con el tema. Se reconocen las contribuciones de otros autores mostrando alguna relación lógica entre sus resultados y la investigación realizada. Se pone énfasis en hallazgos y conclusiones importantes, sin incluir detalles secundarios; además se relacionan esos hallazgos con el diseño y los métodos empleados.

El marco teórico ubica el problema desde una corriente ideológica, desde un sistema de pensamiento. El marco teórico no constituye necesariamente un elemento rotulado con este nombre. Sin embargo el texto de la introducción debe presentar evidencia clara sobre cómo el autor relaciona el problema a investigar con el cuerpo de conocimientos y teorías existentes sobre el tema. Además el marco teórico permite disponer de una gama de opciones posibles de diseño de la investigación. Entendemos el término «diseño» como el enfoque, el abordaje y el plan de la investigación.

La definición operacional es aquella definición en términos de «condiciones observables» u «operaciones» bajo las cuales sucede un evento. Ejemplo: «Buena memoria es el recuerdo exitoso de quince nombres -sobre un total de 20- inmediatamente después de un período de aprendizaje de 5 minutos».

La definición de los atributos o características de los sujetos de la investigación que llamamos «variables» debe ser muy clara y específica.

Estilo

El primer párrafo debe invitar al lector a seguir leyendo. En esta parte tan importante del artículo la organización de ideas debe crear una sensación creciente de descubrimiento. El tono debe ser conversacional, sincero y tener la fuerza de la lógica, no de la insistencia o la exageración. Es preciso ir directamente al grano, liberarse de cualquier elemento que impida la recepción inmediata de las ideas. El escritor debe ponerse en el lugar del lector y evitarle que tenga que volver a leer para entender. Es una buena práctica dar el manuscrito a leer a una persona de confianza que no tema indicarle al autor lo que falta, lo que sobra y lo que no ha quedado claramente expresado.

La introducción debe reescribirse cuantas veces sea necesario después de terminar el manuscrito. Es interesante tener en cuenta que, muchas veces, en la parte de la discusión aparecen frases muy adecuadas para la introducción del trabajo. Debe evitarse ofrecer demasiados datos o datos irrelevantes. (Por ejemplo, la descripción física de una región entera para describir los problemas de una pequeña comunidad). Si bien se recomienda explicar los conceptos difíciles que no estén al alcance de la mayoría de los lectores, debe darse por entendida alguna información obvia, o cuya exactitud no sea necesaria para seguir el curso de la investigación presentada.

A menudo es aconsejable el sacar párrafos enteros del comienzo de la introducción si el escritor tarda demasiado en entrar en materia. El fracaso en delimitar el problema inmediatamente indica vaguedad y confusión del autor sobre lo que desea lograr con el consiguiente fastidio del lector y el peligro de que abandone la lectura del artículo.

Faltas frecuentes

- a. Inexactitud en la identificación del problema general.
- b. Imprecisión en la definición del problema de investigación.
- c. Ausencia de un marco teórico de referencia conceptual.
- d. Insuficiencia de respaldo bibliográfico.
- e. Falta de definición de las variables.
- f. Falta de definición operacional de los términos principales.
- g. Objetivos demasiado generales, vagos o ambiguos.
- h. Hipótesis mal formuladas.
- i. Limitaciones del estudio no explícitas.

5. Materiales y métodos

Concepto y objetivos

Este capítulo es sin duda el más vulnerable para el rechazo de manuscritos de artículos originales por las revistas biomédicas de prestigio que utilizan sistemas de revisión por árbitros, incluyendo entre éstos a algún experto en diseños de investigación y análisis de los datos.

En esta parte se responde a la pregunta: «¿Cómo se realizó la investigación?». Algunos textos sobre investigación científica utilizan otros términos como sinónimos. Tal es el caso de «sujetos y métodos», «pacientes y métodos», «materiales y técnicas», «métodos», «diseño y métodos». ¹⁵

Suele haber estudios aparentemente válidos y potencialmente útiles cuyo informe omite el mínimo de detalles de qué se hizo, a quién se hizo y cómo se hizo. Esto incluye aspectos tan importantes como el procedimiento de muestreo supuestamente aleatorio, la selección de pacientes y de controles, la caracterización de las variables dependientes e independientes, los métodos de análisis, los errores estándares, los intervalos de confianza, etc.

El objetivo de este capítulo es, en consecuencia, describir el diseño de la investigación, incluyendo sus premisas, sus ventajas y sus limitaciones, por una parte, y explicar cómo se llevó a la práctica, justificando la elección de sujetos, de métodos y de técnicas, por la otra..

Surge la pregunta lógica: ¿Cuánto detalle debe incluir esta información?. La respuesta es inequívoca: El detalle suficiente que permita al lector general la clara y recta comprensión del proceso de la investigación. Además el detalle necesario que haga posible reproducir el estudio a otro investigador, adaptándolo a su realidad.

El término *diseño* viene del italiano «disegno» y significa la traza, el delineamiento de una investigación. Es la descripción sucinta o bosquejo del enfoque, de la estrategia de abordaje del problema de investigación para el logro de los objetivos definidos al final de la fase conceptual de la investigación.

El investigador dispone habitualmente de un espectro más o menos amplio de posibles diseños de estudio. En consecuencia debe dar razones para justificar la selección de un diseño en particular y contener el plan de la investigación con los elementos formales del trabajo: fuentes de información, sean estas documentales o de campo; métodos de recolección, tabulación, presentación y análisis de la información.

Todo diseño, más allá de su tipo, debe asegurar de alguna manera dos elementos: a) Control, en el sentido de comprobación o de comparación. De entre las diversas formas de control cabe mencionar el control por manipulación, el control por eliminación o inclusión, el control estadístico y el control por aleatorización; b) Validez, con sus variedades de validez interna y validez externa.¹⁶

Las características de un buen diseño incluyen:

- a. Adecuación a las «preguntas de investigación», formuladas en la forma de objetivos o hipótesis.
- b. Precisión, la cual aumenta cuando disminuye la variabilidad inducida por las variables extrañas.
- c. Poder, entendido como la capacidad de un diseño de investigación para descubrir las relaciones entre variables, para operacionalizar las variables independientes de manera que aumente la capacidad de descubrir diferencias.
- d. Falta de sesgos, es decir, errores que influyen sistemáticamente produciendo una distorsión de los resultados del estudio en alguna dirección.

Los sesgos suelen agruparse en tres tipos: sesgos de selección, sesgos de clasificación y sesgos de confusión. Los dos primeros se deben controlar en la fase de diseño. El último se puede controlar tanto en la fase de diseño como en la de análisis.

La elección del diseño es la decisión más importante de la investigación, después de la del problema de investigación y de la definición de los objetivos de la misma.

El lector espera una clara descripción de los *sujetos* que intervienen en la investigación, aclarando su identidad, su número y el modo en que se seleccionaron. El lector debe obtener una idea muy clara tanto de la muestra como del universo de donde proviene la muestra.

Un aspecto muy importante es la descripción de las *fuentes de información*, tanto documentales como de campo, con algún comentario sobre la confiabilidad de las mismas.

El autor debe explicar la *intervención*, con caracterización de la variable independiente y la variable dependiente. Junto a esta explicación debe señalar los *instrumentos* utilizados tanto para la recolección de la información y para la medición de las variables bajo observación y estudio. Esto incluirá formularios, cuestionarios, equipos, materiales, etc.. empleados en el estudio.

Se espera encontrar una explicación sucinta de los *procedimientos* utilizados en la obtención de los datos y en la tabulación de los mismos con especificación de los criterios de inclusión y exclusión de sujetos a los grupos de estudio y de control.

Una definición de los *términos principales* en forma operacional evitará confusiones innecesarias.

Es crucial el especificar qué *métodos estadísticos* se emplearon para analizar los hallazgos. Esto cobra especial importancia para la determinación de la validez interna (capacidad del investigador de atribuir diferencias en los grupos o participantes a la variable independiente) y de la validez externa (grado en que se pueden generalizar los resultados de la muestra bajo estudio a la población).

Surge aquí nuevamente una pregunta práctica: ¿Cuánta información estadística debe incluir el informe de la investigación redactado en la forma de un artículo original? La cantidad de información estadística y su grado de elaboración dependerá del tema específico a investigar, de los objetivos (o hipótesis) definidos y del tipo de diseño seleccionado.

El Comité Internacional de Editores de Revistas Médicas (CIERM) ha propuesto las siguientes pautas para presentar por escrito los aspectos estadísticos de la investigación:¹¹

- a. Describa los métodos estadísticos con detalles suficientes para que un lector enterado que tenga acceso a los datos originales puede verificar los resultados informados;
- b. Cuantifique los resultados y preséntelos con indicadores apropiados de error o incertidumbre de la medición (por ejemplo, intervalos de confianza);
- c. No dependa exclusivamente de las pruebas de comprobación de hipótesis estadísticas, tal como el uso de los valores P, que no transmiten información cuantitativa importante;
- d. Analice la elegibilidad de los sujetos de experimentación;
- e. Proporcione los detalles del proceso de aleatorización cuando este exista;
- f. Si se usaron medios para encubrir las observaciones (método ciego, doble ciego, etc.) descríbalos junto con los resultados que dieron;
- g. Informe no solamente sobre las bondades de la intervención sino también sobre las complicaciones de la misma (efectos colaterales indeseables);
- h. Especifique el número de observaciones, sus relaciones y posibles correlaciones con otras unidades;
- i. Indique las pérdidas de sujetos de observación (por ejemplo, las personas que abandonan un ensayo clínico);
- j. Utilice referencias actuales sobre el diseño del estudio y los métodos estadísticos (últimos diez años y de preferencia los últimos cinco años), más bien que los artículos originales donde se describieron por primera vez;
- k. Especifique cualquier programa de computación de uso general que haya empleado;

- l. Ubique las descripciones generales de los métodos utilizados en la sección de Materiales y Métodos. Cuando resuma los datos en la sección de Resultados, especifique los métodos estadísticos que se emplearon para analizarlos;
- m. Limite el número de tablas y figuras al mínimo necesario para explicar el tema central del artículo y para evaluar los datos en que se apoya. Aplique el concepto de «lo necesario y lo suficiente»;
- n. Evite el uso no técnico de términos de la estadística, tales como «al azar», «normal», «significativo», «correlación», «muestra», «parámetro»;
- o. Defina los términos, las abreviaturas y la mayor parte de los símbolos estadísticos.

Una debida consideración a cada una de estas pautas ayudará grandemente al autor al momento de preparar y redactar el manuscrito del artículo. También será muy útil en la etapa más temprana del protocolo de investigación, cuando se toman decisiones cruciales acerca de los objetivos y el diseño de la investigación.

Estructura

El formato de este capítulo debe servir fielmente a los objetivos descritos en el cuarto párrafo de «Concepto y Objetivos». De esta manera el texto de material y métodos debe incluir los párrafos correspondientes a: a) Diseño; b) Sujetos; c) Fuentes de información; d) Intervención; e) Instrumentos; f) Procedimientos; g) Términos principales; h) Pruebas (ó modelos) estadísticos.

Estilo

El estilo debe ser narrativo, directo, factual. Deben preferirse frases cortas y sencillas en lugar de largas y complejas; la experiencia muestra que las primeras resultan más fáciles de entender por los lectores. Debe emplearse un estilo objetivo, evitando los pronombres personales. La claridad, la concisión y la secuencia son requisitos esenciales. No tienen cabida cuestiones ni conceptos que no son estrictamente relevantes al tema del estudio y al objetivo del capítulo.

En cuanto a la redacción debe utilizarse el tiempo pasado ya que se trata de relatar lo que sucedió durante la investigación.

Faltas frecuentes

- a. Diseño inapropiado para el objetivo de la investigación.
- b. Diseño apropiado pero en desacuerdo con el nivel actual de conocimientos sobre el problema.

- c. Diseño cuestionable en términos éticos.
- d. Hipótesis mal formuladas.
- e. Falta de definición operacional de los términos principales.
- f. Muestra no representativa del universo
- g. Universo no definido.
- h. Falta descripción de la asignación de sujetos a los grupos de estudio (experimental, control).
- i. Imprecisión en la descripción de los métodos.
- j. Supuestos básicos de la investigación no explícitos.
- k. Falta de explicación de las limitaciones del estudio.
- l. Inclusión anticipada e innecesaria de resultados.
- m. Falta de explicación sobre las pruebas estadísticas aplicadas.
- n. Falta de orden lógico en la argumentación.
- o. Estilo con frases demasiado largas y complicadas.
- p. Mención de asuntos interesantes pero superfluos que distraen y confunden al lector.

6. Resultados

Concepto y objetivos

Los «Resultados» constituyen lo principal, la sustancia, el fondo del artículo científico original en la forma de datos. Todo lo que antecede: la introducción y los materiales y métodos, cobran importancia en la medida que posibilitan la obtención de resultados. Lo que viene a continuación: la discusión para encontrar el significado de los hallazgos y la conclusión con la respuesta a la pregunta de investigación, son la consecuencia directa de los hallazgos debidamente seleccionados, organizados, presentados y analizados.¹⁷

El capítulo de resultados responde a las preguntas: ¿Qué se encontró en la investigación? ¿Cuáles son los hallazgos principales, producto de las metodologías utilizadas? En términos generales, los datos que se recolectan no son directamente compatibles con el análisis por lo cual hay que efectuar una selección, una tabulación, una presentación, una descripción y un análisis de los mismos. El capítulo de «Resultados» -en consecuencia- contiene todos los hallazgos principales, experimentales o de otro tipo, que el autor ha acumulado durante el curso de la investigación, clasificados y ordenados de una manera sistemática.

Surge aquí una pregunta práctica: ¿Con qué grado de detalle técnico deben presentarse los hallazgos? La respuesta es inequívoca: deben aparecer aquellos detalles que se consideren esenciales para una comprensión clara de la información y de su aplicación.

Razones de orden práctico aconsejan el organizar y presentar conjuntos de resultados con valores numéricos en la forma de tablas o de figuras. Si se han recolectado datos extensos, como sería el caso de valores analíticos, lo mejor es simplemente resumirlos en el texto agregando quizás un ejemplo representativo. Otros detalles pueden incluirse en un «Apéndice», si el caso lo requiere.

El investigador debe ser cuidadoso en informar de todos los hallazgos importantes en la forma más exacta y completa posible que apoyen las hipótesis originales del autor. Al mismo tiempo razones de tipo científico y de ética exigen que el investigador no omita aquellos hallazgos negativos significativos. El capítulo de resultados debe también resumir el análisis de los hallazgos utilizando las estadísticas descriptivas en primer lugar, tanto aquellas medidas de tendencia central como aquellas otras de dispersión.

El capítulo de «Resultados» debe presentar solamente hallazgos que se puedan describir con seguridad y con una clara conciencia de que: a) Son confiables; b) Son significativos, y c) Son nuevos. El requisito de confiabilidad se refiere al grado de consistencia o seguridad con la cual un instrumento mide los atributos que intenta medir. El requisito de significación se refiere a la probabilidad de que una relación observada podría deberse al azar y no a la variable independiente. El requisito de novedad se refiere a que los hallazgos aporten algo realmente distinto, en algún sentido, al conocimiento existente sobre el problema al momento de realizarse el estudio, tal como lo muestra una revisión bibliográfica seria, pertinente y actualizada.

Los objetivos del capítulo de resultados pueden enunciarse brevemente como sigue:

- a. Presentar los hallazgos más importantes y pertinentes a los objetivos del estudio en forma comprensible y coherente;
- b. Presentar los hallazgos en un orden lógico, el cual permita una secuencia que apoye la hipótesis;
- c. Mencionar todos los hallazgos relevantes, incluso aquellos contrarios a la hipótesis formulada;
- d. Utilizar el medio de presentación, o combinación de medios, más adecuado, claro y económico (texto, tablas, gráficos);
- e. Presentar los resultados de las pruebas estadísticas aplicadas a la información recolectada;
- f. Describir los hallazgos y su análisis con detalle suficiente como para justificar las conclusiones.

Los resultados resumen la verdadera contribución del estudio realizado. El capítulo de resultados es el referente obligado de la conclusión, la cual se ubica al final de la discusión. La conclusión que no se apoya en los resultados no es legítima.

Estructura

La estructura de este capítulo debe servir los objetivos descritos en el numeral anterior. Para ello se presenta la información seleccionada y ordenada mediante el texto, las tablas y el material de ilustración más pertinentes al tema y a los objetivos del artículo.

El *texto* es la forma principal y más eficiente de presentación de resultados.

El texto debe expresarse en forma comprensible y verificable, tomando en cuenta todas las variables seleccionadas. La redacción sigue una secuencia lógica en relación con los objetivos o las hipótesis a fin de responder a la pregunta de investigación.

Las tablas y los gráficos son auxiliares poderosos del texto. Se debe estar alerta a fin de evitar redundancias.

El texto debe ser claro, preciso y limitarse a lo estrictamente necesario. Un buen texto cumple con requisitos básicos:

- a. Utiliza una sucesión adecuada de párrafos;
- b. Utiliza títulos y/o subtítulos solamente para agregar claridad a la categorización;
- c. Cita todas las tablas y figuras;
- d. Cita todas las referencias bibliográficas;
- e. Se expresa en tiempo pasado cuidando de no repetir lo descrito en materiales y métodos.

Las *tablas* son como la imagen de los datos; consisten en conjuntos ordenados y sistemáticos de números y/o palabras que muestran valores en renglones o filas horizontales y en columnas verticales.

Los objetivos de las tablas son:

- a. Presentar información esencial de tipo repetitivo en una forma fácilmente visible y comprensible.
- b. Mostrar frecuencias, relaciones, contrastes, variaciones y tendencias mediante la presentación ordenada de la información.
- c. Complementar —no duplicar— lo descrito en el texto.

Los requisitos de una buena tabla son:

- a. Ser autoexplicativa.
- b. Ser sencilla y de fácil comprensión.
- c. Tener título breve y claro.
- d. Indicar la fuente de origen de la información (lugar y fecha).
- e. Incluir las unidades de medida en el encabezamiento de cada columna.
- f. Indicar la base de las medidas relativas (razones, porcentajes, tasas, índices, etc.)
- g. Hacer explícitas las abreviaciones.
- h. Indicar las llamadas a notas de pie de tabla mediante letras colocadas como exponentes, en orden alfabético.
- i. Agregar información, no duplicarla.

Las *figuras* son como la imagen de las ideas; consisten en gráficos, diagramas, dibujos lineales, mapas, fotografías, radiografías, etc..

Los objetivos de las ilustraciones son:

- a. Aumentar la información escrita, aclarar conceptos y ofrecer una orientación visual.
- b. Completar el texto sin reemplazarlo, ni requerir leyendas extensas para asegurar su comprensión.
- c. Destacar tendencias.
- d. Hacer comparaciones claras y exactas.

El tipo de información y el objetivo específico determina el tipo de ilustración que ha de utilizarse.

Los requisitos de una buena ilustración son:

- a. Ser autoexplicativa.
- b. Ser sencilla y de fácil comprensión.
- c. Tener título breve y preciso.
- d. Indicar fuente de origen de la información (lugar y fecha).
- e. Indicar claramente las coordenadas, escalas y unidades de medida.
- f. Equilibrar la sencillez con la claridad y la funcionalidad.
- g. Proveer adecuación y uniformidad.
- h. Hacer explícitas las abreviaciones.
- i. Agregar información, no duplicarla.

La decisión sobre la selección del tipo de gráfico es de mucha trascendencia. La naturaleza de los datos obliga a seguir ciertas reglas para asegurar que el gráfico es adecuado al objetivo que se persigue y al tipo de serie que se quiere analizar.

Los gráficos más frecuentemente utilizados en publicaciones biomédicas son:

- a. Gráfico en barras; b. Histograma; c. Polígono de frecuencias; d. Gráfico lineal de escala aritmética; e. Gráfico lineal de escala semi-logarítmica; f. Gráfico en sectores; g. Gráfico de dispersión; h. Diagrama de estrellas; i. Caras de Chernoff.

Se deben seleccionar cuidadosamente las tablas y los gráficos porque son costosos y ocupan mucho espacio en la publicación. Además el exceso de tablas, de material gráfico o de ambos disminuye el efecto que se desea lograr, cual es el de realzar el texto.

Los temas de TABLAS y GRAFICOS se tratan en detalle en los capítulos 7 y 8, respectivamente.

Estilo

La redacción de los resultados estadísticos es —a menudo— una tarea difícil para el investigador por alguna inseguridad sobre qué decir y por la escasa formación y experiencia sobre cual es el mejor estilo para decirlo.¹⁸

El informe de la investigación no debe reclamar que los datos «probaron», «verificaron», «confirmaron» o «demostraron» que la hipótesis era correcta. Las hipótesis «se sustentan» o «no se sustentan»; «se aceptan» o «se rechazan».

Los datos brutos y el detalle de los cálculos estadísticos no tienen cabida en el capítulo de resultados. Por el contrario, el capítulo de resultados debe mostrar los datos recopilados y analizados, utilizando para ello tanto el texto como las tablas y los gráficos.

La información de resultados de pruebas estadísticas requiere incluir tres informaciones complementarias: el valor de la estadística calculada, el número de grados de libertad y el nivel de significación.

Cuando se presentan los resultados de varios análisis, es recomendable resumir los hallazgos en una tabla. Una buena tabla, con título preciso y encabezamientos de columnas claro es una forma importante de evitar frases y párrafos repetitivos aburridos y de ahorrar espacio en la revista. El texto se reserva para destacar aquellos hallazgos más importantes.

Faltas frecuentes

- a. Se incluye información no pertinente a los objetivos del estudio.
- b. Se incluye información pertinente pero de muy escasa importancia para el estudio.
- c. Se confunden hechos (nivel objetivo) con opiniones (nivel subjetivo).
- d. Se presentan hallazgos sin una secuencia lógica.
- e. La información que se incluye no es suficiente para justificar las conclusiones.
- f. El tratamiento estadístico no se menciona, no se explica con claridad, o es inapropiado.
- g. Se utilizan tablas y/o gráficos inadecuados o poco claros.
- h. Las tablas y/o los gráficos son excesivos. De hecho se utilizan de forma indiscriminada, inadecuada e innecesaria. Hay duplicación y redundancia en la presentación de la información.
- i. Los gráficos son inadecuados para destacar tendencias y hacer comparaciones.
- j. Los gráficos incluyen muchas series estadísticas juntas y/o éstas tienen muchas categorías, lo cual complica el dibujo y dificulta su comprensión e interpretación.
- k. Los gráficos se presentan incompletos. Falta información sobre número, título, escalas, fuentes y simbología.
- l. Las tablas y los gráficos carecen de legibilidad y de comprensibilidad.

7. Tablas

La Sección “3.1. Tablas” del Capítulo A. “Análisis descriptivo”, de la Parte II de esta obra (páginas 133-136) introduce el tema de las tablas como una forma de presentar datos de tipo numérico en forma organizada y resumida. El texto de este acápite relaciona las tablas con los conceptos de frecuencia absoluta, frecuencia relativa, frecuencia acumulativa absoluta, frecuencia acumulativa relativa e intervalos de valores.

Introducción

El análisis de los datos es un componente muy importante en la práctica de la investigación clínica y epidemiológica. Para analizar los datos en forma eficiente el investigador debe primero familiarizarse con los datos antes de aplicar técnicas analíticas. De esta manera el investigador que empezó examinando registros individuales avanzará organizando y resumiendo los datos en tablas apropiadas.

Algunas veces las tablas resultantes constituyen el único elemento de análisis necesario, especialmente cuando la cantidad de datos es pequeña y las relaciones son sencillas.

Cuando los datos son más complejos será necesario recurrir a gráficos y/o a esquemas para visualizar patrones más amplios y tendencias, como así mismo para identificar variaciones y prospecciones desde esas tendencias. Estas variaciones pueden representar hallazgos nuevos importantes o simplemente detectar errores de codificación o de transcripción.

Una vez que se ha completado el análisis, las tablas, los gráficos y los esquemas sirven como ayudas visuales útiles para describir los datos a otras personas. De lo anterior se desprende que los objetivos de las tablas, de los gráficos y de los esquemas son: a) posibilitar la comprensión y el análisis de los datos; b) comunicar información acerca de los datos.

Las tablas organizan y condensan datos en una forma estandarizada; se construyen para presentar información en forma numérica, especialmente de tipo repetitivo: números de ocurrencia, porcentajes, valores en escalas y otras características de los fenómenos bajo estudio. Una buena tabla es un conjunto ordenado y sistemático de números y/o palabras que muestra valores en renglones o filas (horizontales) y en columnas (verticales).

La facilidad y la velocidad para comprender una tabla depende mucho de la disposición lógica de su tabulación. Los encabezamientos de filas y de columnas se relacionan con el conocimiento que posee el lector; la información contenida en las diversas casillas del cuerpo de la tabla corresponde al conocimiento que busca el lector y que ofrece el autor.

La estructura de la tabla debe ser lo más simple y clara posible. Las presentaciones tabulares unidimensionales no suelen presentar problemas para el lector promedio. Sin embargo algunos lectores experimentan dificultades para comprender una tabla tridimensional o matriz de triple vía. Nunca se ha de olvidar la importancia de presentar la información en una secuencia que sea compatible con la estrategia de procesamiento de información del lector. Esta estrategia opera usualmente de izquierda a derecha y de arriba hacia abajo.¹⁹

Una tabla permite disponer y ordenar grandes cantidades de datos en columnas y en filas de tal forma que el lector pueda observarlos, estudiarlos, comprenderlos y analizarlos tanto en forma individual como en conjunto. Una tabla bien diseñada suele ser más exacta, más completa y puede presentar mucha más información que un gráfico. Las tablas están particularmente indicadas para el manejo de información com-

pleja en escritura científica, en informes de progreso de proyectos, en informes estadísticos, en informes económicos y de negocios, etc. Una buena tabla condensa información en una forma fácil y rápidamente comprensible y la ubica en un espacio limitado; cuando está bien diseñada ahorra ventajosamente muchos párrafos y hasta páginas de texto expositivo.

Las tablas —para ser útiles— deben ser exactas, fáciles de leer y de comprender. Esto requiere título, espaciamiento, rayado, arreglo de los encabezamientos y ubicación de la tabla con respecto al texto. Las tablas completan el texto, razón por la cual deben ubicarse lo más cerca posible del párrafo donde se las menciona.

En cualquier tabla, el centro de atención consiste en una categoría de respuestas. La tabla se construye para ilustrar las variaciones (variables dependientes) con respecto a esa categoría de respuestas (variables independientes). Un ejemplo simple sería el de un estudio sobre el hábito de fumar cigarrillos en los adolescentes en España. Aquí la variable dependiente sería el hecho de si el respondente fuma o no fuma. Las variables intervinientes serían cualesquiera de las categorías o hechos relacionados con el respondente y seleccionados por el investigador: edad, sexo, escolaridad, hábito de fumar de los padres, ingreso económico, exposición a campaña antitabáquica, etc.

La inclusión de información de más de dos variables en la tabla obliga a subdividirla, a expandirla y la hace más complicada hasta llegar a un límite donde se vulneran los requisitos de claridad y de eficiencia. La solución aquí está en replantearse la importancia real de cada dato; si el investigador le otorga mérito a todas las variables estudiadas, entonces podría justificarse el agregar otra tabla.²⁰

Diseño y construcción de tablas

Hoy existen muchas convenciones para construir tablas. Tanto las asociaciones científicas y profesionales como algunos organismos internacionales especifican sus convenciones en sus respectivos manuales de estilo o de publicaciones. Algunos de estos manuales son de tipo restrictivo para uso interno; otros, en cambio, se publican para una amplia distribución a un público interesado compuesto de investigadores, profesores, directores de revistas, correctores de estilo y casas editoriales. J.B. Howell describe en su libro ya clásico 231 manuales de estilo en uso en países de habla inglesa agrupados en 27 temas, biología y ciencias entre otros.²¹

La comunidad científica internacional ha alcanzado, en general, consenso sobre los fundamentos o bases para la presentación de información en forma tabular. Esto incluye elementos importantes tales como: numeración, título, encabezamientos de

columnas, cuerpo, intervalos de clases, tamaño y forma, abreviaturas y símbolos, alineación, omisiones, notas a pie de tabla, fuentes de información y ubicación con respecto al texto.²²

1. Numeración

Toda tabla debe ir numerada. Hay dos sistemas de numeración de tablas: a) numeración correlativa en orden de mención en el texto completo; b) numeración por capítulos según secuencia de aparición en cada capítulo. En el primer caso la tabla llevará un número simple; en el segundo caso llevará un número compuesto por el número del capítulo, primero, y el número de orden de la tabla en ese capítulo, segundo. A modo de ejemplo, si la primera tabla aparece en el capítulo 3, en el primer caso llevará el número 1; en el segundo caso llevará el número 3.1.

Se debe asignar un número árabe, consecutivo a cada tabla aun cuando haya pocas tablas en el documento. El orden en el cual se mencionan las tablas en el texto determina la numeración. Las tablas incluidas en los apéndices deben numerarse consecutivamente con las tablas en el texto. Esto significa que si la última tabla en el texto es la tabla 7, la primera tabla del apéndice será la tabla 8. Toda referencia en el texto a una tabla debe ser por número; en consecuencia no se debe usar una frase introductoria tal como «En la tabla siguiente se muestra...» sino «La tabla 1 muestra...». Las referencias numeradas facilitan al lector el ubicar la tabla sin demora, aun cuando —por razones de diseño— la tabla no se ubique inmediatamente después del párrafo donde se la cita.

2. Título

Toda tabla, además de ir numerada en el margen superior izquierdo, debe tener un título. El título debe describir el contenido de la tabla en forma breve y exacta. En general se recomienda un máximo de 10 palabras. No es aceptable el utilizar títulos ambiguos, efectistas, llamativos, no convencionales. El título debe emplear palabras completas, no abreviaturas o siglas, salvo casos especiales.

El título debe escribirse en la misma línea del número de la tabla, separado por un guión y un espacio. El título debe ocupar líneas completas equivalentes al ancho total de la tabla y centrando la línea última si ésta resulta ser más corta. El título debe ir en mayúsculas, debiendo utilizarse igual sistema de mayúsculas en los títulos de todas las tablas. El título debe identificar brevemente el contenido de la tabla sin adelantar información sobre los resultados; la tabla debe mostrar meramente los hechos, reservando para el texto la discusión y la interpretación de esos hechos.

Excepcionalmente el título puede llevar un subtítulo aclarativo. Tal suele ser el caso de especificaciones como «En millones de pesetas» o «N=253»; en ambas situaciones la indicación aplica a toda la tabla.

3. Encabezamientos de columnas

Las columnas de una tabla precisan de encabezamientos en la parte superior, los cuales consisten en indicaciones muy sumarias y precisas del material que contiene la columna. Las cabezas de las columnas pueden incluir subencabezamientos cuando sea necesario, los cuales pueden ser abreviaturas y se imprimen entre paréntesis. El ancho de cada encabezamiento de columna y el número de columnas determinan el ancho total de la tabla, razón por la cual los encabezamientos deben redactarse lo más breve posible.

Los encabezamientos deben ser muy específicos ya que su objetivo es identificar exactamente la información contenida en la columna o grupo de columnas. En el caso del encabezamiento de la primera columna o «columna talón», ésta se ubica en el extremo superior izquierdo de la tabla e identifica una serie de variables, objetos, personas, sucesos acerca de lo cual la tabla presenta información.

4. Cuerpo

El cuerpo de la tabla debe contener todos los datos que ilustran el título de la misma. La consistencia aquí es muy importante; si los datos se informan en palabras, hay que ser consistente con la forma, con la puntuación y con las abreviaturas. Si los datos se informan en números, hay que expresarlos en las mismas unidades. Los números enteros deben alinearse por la columna derecha; los números decimales deben alinearse por sus puntos decimales.

5. Intervalos de clases

La mayoría de las variables de tipo discreto tienen un número limitado de respuestas posibles. Estas respuestas proveen categorías convenientes para usar en una tabla. Tal es el caso de sexo, paridad, grupo sanguíneo, nivel educativo, grado de malignidad de un tumor (estadio), número de sesiones de tratamiento quimioterápico, distritos de una población, drogas a comparar en un tratamiento, tiempo de supervivencia, hábito tabáquico, exposición a un riesgo, etc.

Cuando se trata de variables de tipo continuo tales como tiempo, edad, dosis de irradiación, ingreso económico, presión arterial, etc., éstas tienen un rango amplio de

repuestas posibles. En consecuencia se deben agrupar en un número manejable de categorías, las cuales se denominan *intervalos de clase*.

Para crear los intervalos de clase se deben tener presente las siguientes pautas directrices: a) crear intervalos de clase que sean mutuamente excluyentes; b) usar un número relativamente amplio de intervalos de clase más bien angostos para el análisis inicial; más tarde se pueden combinar los intervalos; c) usar intervalos de clase con significación natural o biológica; d) utilizar agrupaciones por edad de tipo estándar y/o de uso frecuente en la especialidad; e) utilizar los mismos intervalos para el numerador que para el denominador en caso que se vaya a calcular tasas; f) crear categorías para “desconocido”, “no declara”, “no consta”, etc.

En el caso de que no existan intervalos de clase naturales, o de que no haya estándares, se deben usar algunas de las siguientes estrategias para crear intervalos de clase: a) dividir los datos en grupos de tamaño similar; b) basar los intervalos en promedios y desviaciones estándar; c) dividir el rango en intervalos de clase igual.

6. Ejemplos de tablas con intervalos de clases

Tabla IV.1.- Distribución de la duración del tratamiento en 200 pacientes sometidos a litotricia.

Duración (en minutos)	Frecuencia absoluta	Marca de clase	Frecuencia acumulativa absoluta	Frecuencia Acumulativa relativa
$10 \leq x < 20$	4	15	4	0.0200
$20 \leq x < 30$	5	25	9	0.0450
$30 \leq x < 40$	21	35	30	0.1500
$40 \leq x < 50$	27	45	57	0.2850
$50 \leq x < 60$	48	55	105	0.5250
$60 \leq x < 70$	53	65	158	0.7900
$70 \leq x < 80$	28	75	186	0.9300
$80 \leq x < 90$	13	85	199	0.9950
$90 \leq x < 100$	1	95	200	1.0000
Total	200	1.0000		

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.1 muestra la variable independiente, duración de la litotricia en minutos, dividida en nueve intervalos de clase.

Tabla IV.2.- Incidencia de leucemia entre los residentes de Hiroshima, de 1945 a 1966.

DOSIS IRRADIACIÓN	Nº PERSONAS EXPUESTAS	AÑO-PERSONA OBSERVACIÓN	CASOS DE LEUCEMIA
< 5	62.515	915.000	27
5 - 19	10.541	156.000	8
20 - 49	4.555	67.000	14
50 - 99	2.611	38.000	7
100 - 199	1.652	24.100	10
200 - 299	606	9.000	5
300 y más	825	12.100	17
TOTAL	83.305	1.221.200	88

Fuente: Martínez Oliva L. Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos. México: Centro Panamericano de Ecología y Salud, 1990.

La Tabla IV.2 presenta la variable independiente, dosis de irradiación en RADS, dividida en siete intervalos de clase.

Tabla IV.3.- Distribución de frecuencias de niveles de seroalbúmina

Los valores de seroalbúmina en g/litro de sangre de 50 mujeres vistas en una encuesta son los siguientes:	Seroalbúmina (g./litro)	Nº de observaciones
42, 41, 42, 44, 44, 36, 38, 41, 42, 44	30-33	3
42, 39, 49, 40, 45, 32, 34, 43, 37, 39	34-37	7
41, 39, 48, 42, 43, 33, 43, 35, 32, 34	38-41	14
39, 35, 43, 44, 47, 40, 39, 42, 41, 46	42-45	17
37, 49, 41, 39, 43, 42, 47, 48, 51, 52	46-49	7
Utilizando seis intervalos iguales, estos datos se presentan en la Tabla 3 de distribución de frecuencias.	50-53	2
Los rasgos evidentes comprenden ahora la distribución de los valores de seroalbúmina entre las 50 personas estudiadas, su variación desde 30 hasta 53 y el hecho de que un número apreciable de personas tienen niveles de seroalbúmina entre 38 y 45 g/litro.	Total	50

Fuente: Lwanga SK, Tye CY. La enseñanza de la estadística sanitaria. Ginebra, 1987.

La Tabla IV.3 muestra la variable bajo estudio, distribución de frecuencias de niveles de seroalbúmina en g/litro de sangre, dividida en seis intervalos iguales de 4 gramos de seroalbúmina por litro (rango es 20).

Tabla IV.4.- Número anual de consultas médicas por edad en EE.UU.

EDAD	Consultas médicas anuales
menos de 6	6,7
6-16	3,3
17-24	4,4
25-44	4,8
45-64	6,4
65 y más	8,9

Fuente: Ministerio de Comercio de Estados Unidos. Resumen estadístico de la 110ª edición. Washington, D.C.: Oficina del Censo, 1990.

La Tabla IV.4 muestra el número anual de consultas médicas en Estados Unidos por edad, dividido en seis intervalos de clase.

Tabla IV.5.- Casos nuevos reportados de sífilis primaria y secundaria. Tasas específicas por edad y raza por 100.000 habitantes. EE.UU., 1989.

Grupo de edad (años)	Tasa según raza (por 100.000)			Total
	Blancos	Negros	Otros	
≤14	0,0	2,4	0,8	0,4
15-19	2,4	131,5	51,0	24,3
20-24	5,8	323,0	139,2	55,9
25-29	5,4	270,9	117,9	44,1
30-34	4,7	256,6	83,2	38,8
35-44	2,9	135,0	47,8	19,0
45-54	1,7	76,7	29,6	10,5
≥55	0,5	19,4	10,4	2,4
Total	2,2	115,8	45,8	17,7

Fuente: Centro para el Control de Enfermedades. Informe de toxi-infecciones alimenticias. Atlanta: CCE, 1989.

La Tabla IV.5 presenta la variable edad dividida en ocho grupos con intervalos desiguales: a) el primer grupo con intervalo de menos de 14 años; b) los siguientes cuatro grupos con intervalos de 5 años; c) los siguientes dos grupos con intervalos de 10 años; d) el último grupo con intervalo mayor de 55 años.

7. Tamaño y forma

En la mayoría de las tablas las columnas corren el camino largo de la tabla. Si la tabla es larga y angosta se puede ahorrar espacio mediante un diseño en «dobladura» o «pliegue», dividiéndola en partes iguales y colocándolas una al lado de la otra, separadas mediante una doble línea vertical.

Si la tabla es demasiado ancha para la página, debe ubicarse en forma transversal (de lado) a lo largo de la página; en este caso el número de la página debe aparecer en el lugar habitual y no debe aparecer ningún texto en la página que contiene la tabla de lado.

Si la tabla no pudiera acomodarse de lado por ser demasiado ancha, entonces se procede a ubicarla en dos páginas cara a cara, asegurándose que las partes de la tabla tengan las mismas dimensiones en ambas páginas y que la encuadernación asegure una exacta concordancia de las líneas horizontales. Otra solución a la ubicación de una tabla demasiado ancha es la de imprimirla en una hoja especial de tamaño mayor e insertarla entre dos páginas con las dobles correspondientes.

Si la tabla es muy larga, puede continuarse de una página a las páginas siguientes. El número de la tabla junto con el título se incluye en el comienzo de la tabla; las páginas siguientes llevarán solamente el número de la tabla con la indicación «Continuación»; sin embargo los encabezamientos de columnas se deben repetir en cada página.

8. Abreviaturas y símbolos

Aunque su uso está proscrito en los títulos —excepto para símbolos matemáticos, químicos y biológicos de uso común— las abreviaturas y los símbolos constituyen ahorro legítimo de espacio en encabezamientos de columnas y en el cuerpo de la tabla. Se debe utilizar abreviaciones estándar que sean autoexplicativas; en caso contrario debe incluirse una nota explicativa a pie de tabla, en el ángulo inferior izquierdo.

9. Alineación

En una columna de cifras se deben alinear todos los puntos decimales y las comas. Igualmente se deben alinear signos monetarios, signos más, signos menos, signo mayor que, signo menor que, signo igual, etc.

10. Omisiones

En una columna larga de cifras, el número cero (0) que precede a un punto decimal puede suprimirse en todas las entradas excepto la primera y la última. Si todas las cifras están en miles o millones, se puede ahorrar espacio suprimiendo los ceros relevantes y anotando el hecho al final del título, como por ejemplo: «(cifras en millones)».

11. Notas a pie de tabla

Las notas a pie de tabla se utilizan para explicar cualquier parte de la tabla que lo requiera. Las notas a pie de tabla son de cuatro tipos: a) notas sobre fuente de los

datos; b) notas generales; c) notas sobre partes específicas de la tabla; d) notas sobre nivel de probabilidad.

- a. Si los datos de una tabla no provienen del estudio que se presenta sino que se han tomado de otra fuente, el autor debe señalarlo claramente en la primera nota a pie de tabla, la cual se inicia con la palabra «Fuente:».
- b. Las notas generales incluyen comentarios sobre la validez de los datos que se presentan, sobre cómo se obtuvo la información, etc.
- c. Las notas sobre partes específicas incluyen explicaciones sucintas sobre terminología, sobre unidades utilizadas o sobre algún otro aspecto relevante de la tabla.
- d. Cuando la tabla contiene valores con niveles de probabilidad conocida ésta se debe incluir en una nota a pie de tabla, las cuales se indican con asteriscos. Se utiliza un asterisco para el nivel más bajo de probabilidad, dos asteriscos para el siguiente y así sucesivamente hasta el más alto.

Las notas a pie de tabla no se deben numerar en la misma serie que las notas a pie de página del texto.

12. Fuente de información

La línea de «Fuente» da crédito a publicaciones, a personas o instituciones que proveyeron la información utilizada para crear la tabla o la tabla misma. Se deben ubicar en el extremo inferior izquierdo de la tabla y debe contener el texto de origen, el nombre propio de la persona o de la institución y la fecha de su confección.

13. Ubicación

Las tablas deben ir integradas en el texto; excepcionalmente pueden insertarse al final como un apéndice. La audiencia general aprecia más las tablas en el texto; en este caso el autor debe preparar al lector anticipando la presencia y discusión de la tabla en el texto, en el párrafo anterior a la tabla misma.

Idealmente cada tabla deberá ubicarse lo más cerca posible del final del párrafo donde se la cita por primera vez. Si la tabla no puede acomodarse en el espacio restante de la página es preferible continuar con el texto para completar la página y ubicar la tabla al comienzo de la página siguiente. Si la tabla aparece intercalada en el texto, deben dejarse tres espacios en blanco encima y tres espacios en blanco debajo de cada tabla.

Una audiencia de expertos podría interesarse en el conjunto de tablas en un apéndice; en este caso conviene organizar las tablas según algún orden lógico. También debe plantearse la conveniencia de agregar alguna tabla-resumen.

Una buena tabla es autoexplicativa de manera que si la desprendemos del texto, la información que presenta es autosuficiente para su comprensión. Una buena tabla es limpia, es clara y es legible. Con demasiada frecuencia los escritores principiantes tratan de acomodar demasiados datos en la misma tabla; en efecto esta se presenta con numerosos encabezamientos y subencabezamientos de columnas, lo cual atenta contra el principio fundamental de claridad de la tabla.¹²

14. Tablas de una variable

Las tablas de una variable son las tablas más básicas en epidemiología descriptiva; su objetivo es mostrar distribución de frecuencia de una sola variable, ej. grado de malignidad del tumor, número de sesiones de tratamiento, etc.

15. Ejemplos de tablas de una variable

Tabla IV.6.- Distribución del grado de malignidad del tumor en 74 individuos con cáncer de próstata.

GLEASON	Frecuencia absoluta	Frecuencia relativa
Grado 4	12	0.1622
Grado 5	18	0.2432
Grado 6	21	0.2838
Grado 7	14	0.1892
Grado 8 o más	9	0.1216
Total	74	1.0000

Fuente: Burgos R, Chicharro JA, Bobenieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.6 muestra tres encabezamientos de columnas: variable de Gleason (1ª columna); frecuencia absoluta (2ª columna) y frecuencia relativa (3ª columna). Las filas de la 1ª columna contiene las categorías o valores de la variable “grado de malignidad del tumor”. Las filas de la 2ª y 3ª columna contienen el número de veces que se repite cada categoría (frecuencia absoluta y frecuencia relativa, respectivamente). La última fila contiene los totales de ambas frecuencias.

Tabla IV.7.- Distribución del número de sesiones de tratamiento en 200 pacientes sometidos a tratamiento mediante litotricia.

Nº de sesiones	Frecuencia absoluta	Frecuencia relativa
1	165	0.825
2	26	0.130
3	7	0.035
4 o más	2	0.010
Total	200	1.000

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.7 es similar en su estructura a la Tabla IV.4, ya que ambas presentan variables cuantitativas de tipo discreto. La diferencia está en que en la Tabla IV.4 la variable es el grado de malignidad del tumor y en la Tabla IV.5 la variable es el número de sesiones de tratamiento a los pacientes.

16. Tablas de dos variables

Los datos de un estudio pueden también tabularse en forma cruzada para mostrar el recuento de una segunda variable, atributo o característica.

17. Ejemplos de tablas de dos variables

Tabla IV.8.- Trasplantes de riñón y de corazón en Andalucía 1986-1993.

Año	1986	1987	1988	1989	1990	1991	1992	1993
Trasplantes								
Riñón:	183	144	144	124	153	161	206	187
Corazón:	3	3	10	12	24	43	50	51
TOTAL	186	147	154	136	177	204	256	238

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.8 para datos bidimensionales muestra el número de trasplantes de riñón y de corazón, respectivamente, en Andalucía en el período 1986-1993. Esta tabla mide dos variables: a) trasplante (riñón o corazón); b) año del trasplante.

La misma información de la Tabla IV.8 se presenta en el gráfico IV.23, el cual permite observar las fluctuaciones en el número de trasplantes renales y cardíacos en Andalucía en 1986-1993.

Tabla IV.9.- Tipos de acomodación de personas mayores de 65 años en EE.UU.

	TOTAL	HOMBRES	MUJERES
Viven solos	30,4	16,2	40,6
Viven con el cónyuge	54,5	75,1	39,9
Viven con otro pariente	12,8	6,7	17,2
Viven con alguien que no es pariente	2,3	2,1	2,4

Fuente: Ministerio de Comercio de Estados Unidos. Resumen estadístico de la 110ª edición. Washington, D.C.: Oficina del Censo, 1990.

La Tabla IV.9 para datos bidimensionales muestra los tipos de acomodación de la gente mayor de 65 años en Estados Unidos, por sexo.

La Tabla IV.4 (p.357) para datos bidimensionales muestra el número anual de consultas médicas en Estados Unidos, por grupos de edad.

18. Tablas de contingencia o tablas de dos por dos (tablas cuadrículas)

Se trata de un tipo especial de tabla de dos variables con tabulación cruzada de datos en que cada una de las dos variables tiene dos categorías. Las tablas de contingencia se usan frecuentemente para mostrar datos, calcular medidas de asociación y pruebas de significación estadística. También se utilizan en epidemiología para estudiar la asociación entre una exposición y una enfermedad, para comparar personas con y sin exposición, y cada una de aquellas con y sin enfermedad.

19. Ejemplos de tablas de contingencia

Tabla IV.10.- Formato general de una tabla de dos por dos.

	ENFERMOS	SANOS	TOTAL
EXPUESTOS	<i>a</i>	<i>b</i>	Horizontal 1
NO EXPUESTOS	<i>c</i>	<i>d</i>	Horizontal 2
	Vertical 1	Vertical 2	

La Tabla IV.10 muestra el formato genérico de las tablas de dos por dos. El *status* del efecto (enfermo/sano) se indica habitualmente a lo largo de la parte superior de la tabla (encabezamientos de columna); el *status* de la exposición (expuesto/no expuesto) se indica a lo largo de la parte lateral izquierda de la tabla (encabezamientos de fila). Las letras a, b, c y d dentro de las cuatro celdas de la tabla de dos por dos se refiere al número de personas con el *status* de efecto indicado arriba y el *status* de exposición indicado a la izquierda.

Tabla IV.11.- Disposición de los datos en un estudio de prevalencia.

Variable independiente	Variable dependiente			
	Presente	Ausente	Total	Tasa
Presente	a	b	m_1	a/m_1
Ausente	c	d	m_2	c/m_2
Total	n_1	n_2	n	n/n_1

Fuente: Martínez Oliva L. Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos. México: Centro Panamericano de Ecología y Salud, 1990.

La Tabla IV.11 muestra la aplicación de la tabla de dos por dos a la disposición de un estudio de prevalencia con indicación de las variable independiente y dependiente, los totales y las tasas.

Tabla IV.12.- Medidas de riesgo en estudio de cohortes para fumadores respecto a cáncer de pulmón e infarto de miocardio.

	Población	Tasas anuales x 100.000	
		Cáncer pulmonar	Infarto agudo de miocardio
Fumadores	400.000	166,0	599,0
No fumadores	600.000	7,0	422,0
Total	1.000.000	70,6	492,8

Fuente: Martínez Oliva L. Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos. México: Centro Panamericano de Ecología y Salud, 1990.

La Tabla IV.12 muestra la aplicación de una tabla de dos por dos a medidas de riesgo en estudio de cohortes.

Tabla IV.13.- Resultados de una prueba de tamizaje con relación a un examen diagnóstico.

		Resultado del examen diagnóstico	
		Enfermo	No enfermo
Resultado del tamizaje	Positivo	Verdadero positivo (VP)	Falso positivo (FP)
	Negativo	Falso negativo (FN)	Verdadero negativo (VN)

Fuente: Martínez Oliva L. Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos. México: Centro Panamericano de Ecología y Salud, 1990.

Los resultados posibles son:

Verdaderos positivos (VP): individuos que resultaron positivos al tamizaje y que realmente tienen la enfermedad.

Falsos negativos (FN): individuos que respondieron negativamente al examen y que padecen realmente la enfermedad.

Falsos positivos (FP): individuos que resultaron positivos al tamizaje pero en los cuales se descartó totalmente la enfermedad.

Verdaderos negativos (VN): individuos en que el tamizaje fue negativo y realmente no tienen la enfermedad.

Como puede apreciarse, el examen de tamizaje puede dar resultados falsos, de ahí que para evaluar la validez de una prueba se recurra a establecer su grado de sensibilidad y de especificidad.

La Tabla IV.13 presenta el armazón (formato) de una prueba de tamizaje en relación con un examen diagnóstico.

20. Tablas de tres o más variables

Algunas veces se requiere incluir una tercera variable para mostrar un conjunto de datos y sus relaciones en forma más completa. Una tabla de más de tres variables resulta -en general- un tanto congestionada y en la medida que aumenta la complejidad de la tabla suele disminuir su claridad. La decisión de construir una tabla de tres o más variables debe tomarse frente a la conveniencia de agregar una tabla adicional con una o dos variables.

21. Ejemplos de tablas de tres a más variables

Tabla IV.14.- Prestaciones farmacéuticas en los distritos sanitarios de Granada. 1992.

Distrito	Población	Gasto total (en miles)	Gasto por habitante
Distrito 1	73423	944954	12870
Distrito 2	57110	702510	12301
Distrito 3	91771	1341967	14623
Distrito 4	40479	692393	17105
Distrito 5	202758	3128556	15430
Distrito 6	65052	972983	14957
Distrito 7	217772	3255474	14949
Distrito 8	74937	1229192	16403
Total	823302	12268029	118638

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.14 muestra tres variables relacionadas: gasto farmacéutico, población y distrito sanitario.

Tabla IV.15.- Indicadores en seis grandes hospitales andaluces. 1992.

Hospital	Estancia media	Índice de ocupación	Utilización de quirófano	Lista de espera	Mortalidad
Hospital 1	10.60	87.25	57.88	128.15	35
Hospital 2	8.91	74.58	69.99	196.10	26
Hospital 3	11.01	89.27	66.33	65.15	31
Hospital 4	11.19	83.22	70.70	134.87	26
Hospital 5	8.95	85.46	66.77	113.50	28
Hospital 6	9.06	72.08	65.58	66.21	34

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. Metodología de investigación y escritura científica en clínica. Granada: Escuela Andaluza de Salud Pública, 1994.

La Tabla IV.15 muestra una serie de cinco variables (indicadores): estancia media, índice de ocupación, utilización de quirófano, lista de espera y mortalidad, los cuales se relacionan con la gestión de seis grandes hospitales andaluces.

Tabla IV.16.- Tasas de ataque por ítems servidos en la cena de la iglesia.

	Nº de personas que comieron ítems específicos				Nº de personas que no comieron ítems específicos			
	Enfermo	Sano	Total	Tasa de ataque(%)	enfermo	Sano	Total	Tasa de ataque(%)
Jamón asado	29	17	46	63	17	12	29	59
Espinacas	26	17	43	60	20	12	32	62
Puré de patatas	23	14	37	62	23	14	37	62
Ensalada de repollos	18	10	28	64	28	19	47	60
Gelatina	16	7	23	70	30	22	52	58
Croissants	21	16	37	57	25	13	38	66
Pan integral	18	9	27	67	28	20	48	58
Leche	2	2	4	50	44	27	71	62
Café	19	12	31	61	27	17	44	61
Agua	13	11	24	80	3	18	21	14
Pasteles	27	13	40	67	19	16	35	54
Helado de vainilla	43	11	54	80	3	18	21	14
Helado de chocolate	25	22	47	53	20	7	27	74
Macedonia de frutas	4	2	6	67	42	27	69	61

Fuente: Centro para el Control de Enfermedades. Informe de toxi-infecciones alimenticias. Atlanta: CCE, 1989.

La Tabla IV.16. muestra una investigación por toxi-infección alimenticia con tres variables relacionadas: tipo de alimento; condición de haber comido o no haber comido cada alimento; condición de salud (enfermo/sano).

La Tabla IV.5 (p.357) muestra casos nuevos reportados de sífilis, tasas por 100.000, según grupos de edad y raza.

22. Tablas con otras medidas estadísticas

Los diversos tipos de tablas descritos muestran recuentos de casos. Las celdas de una tabla también pueden contener promedios, desviación típica, mediana, tasas, años perdidos de vida potencial, riesgos relativos, intervalos de confianza, etc.

23. Ejemplos de tablas con otras medidas estadísticas

Tabla IV.17.- Media, desviación típica y mediana de los ítems y la escala global del cuestionario de apoyo social. Centro de Salud de Cartuja, 1989.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	Escala
\bar{X} :	2,9	2,7	3,0	3,6	3,7	3,3	3,3	3,2	2,5	3,3	3,6	35,55
S:	1,4	1,3	1,3	1,2	1,3	1,3	1,3	1,2	1,4	1,2	1,3	10,23
Me:	3	3	3	4	4	3	4	3	2	3	4	36

Fuente: De La Revilla L, Bailón E, Luna J, Delgado A, Prados MA, Fleitas L. Validación de una escala de apoyo social funcional para su uso en la consulta del médico de familia. *Atención Primaria* 1991; 8: (9): 52.

La Tabla IV.17 expone los valores de la media, la desviación típica y la mediana de los ítems y la escala global de un cuestionario de apoyo social.

Tabla IV.18.- Mortalidad por tétanos neonatal en el área rural de Jalisco, resultados del análisis simple, 1988.

Variables	Defunciones por			IC(95%)	p ^f
	TN	Otras causas	RM		
Antecedentes de hijos muertos					
2 ó más	4	4	7,00	(0,94-55,55)	0,036
1 ó menos	4	28			
Nivel de hacinamiento					
Hacinado	7	15	7,93	(0,78-182,58)	0,044
Semihacinado	1	17			
Escolaridad de la madre					
Analfabeta	5	6	7,22	(1,05-55,55)	0,024
Primaria incompleta	3	26			
Lugar del parto					
Casa	7	9	17,89	(1,71-447,22)	0,003

Fuente: Tapia Conyer R, Sepúlveda Amor J, Salvatierra Izaba B. Factores asociados a la mortalidad por tétanos neonatal en el área rural de Jalisco. *Salud Pública de México* 1991; 33(3): 212.

La Tabla IV-18 expone la razón de mortalidad, intervalo de confianza y probabilidad exacta de Fisher para cuatro variables relacionadas con la mortalidad por tétanos neonatal.

24. Matriz (formato o plantilla) de tablas

Si bien es cierto que no podemos analizar datos antes de haberlos recolectado, no es menos cierto que deberíamos diseñar la matriz de las tablas por adelantado como parte del plan de su análisis, una vez recolectados los datos. De hecho la mayoría de los protocolos de investigación -los cuales obviamente se preparan antes de que se realice el estudio- requieren de una explicación sobre cómo se analizarán los datos.

Como parte del plan de análisis debemos desarrollar la *matriz o formato de tablas* los cuales muestran cómo se van a organizar y a exponer los datos. Las matrices de tablas son tablas completas, excepto que carecen de los datos. En consecuencia deben mostrar título, encabezamientos de columnas, intervalos de clase, categorías, medidas estadísticas a calcular, totales, etc.

25. Ejemplos de matriz o formato de tablas

Los siguientes formatos de matrices de tabla se diseñaron antes de realizar un estudio de casos y controles de la enfermedad de Kawasaki o síndrome de nódulos linfáticos mucocutáneos febril agudo, enfermedad de los niños de etiología desconocida, la cual ocurre ocasionalmente en grupos. Se decidió probar dos hipótesis -mediante un estudio de casos y controles-: a) asociación del síndrome con enfermedad viral anterior, y b) asociación del síndrome con una exposición reciente a un *champú de alfombra*.

Matriz de tabla IV.19.- Características clínicas de los casos de síndrome de Kawasaki con brote en Octubre-Diciembre, 1994.

Características clínicas	Nº con característica	Porcentaje
Fiebre \geq 5 días	—	()
Inyección conjuntiva bilateral	—	()
Cambios orales		
Labios inyectados	—	()
Faringe inyectada	—	()
Labios secos y partidos	—	()
Lengua en fresa	—	()
Cambios en las extremidades periféricas		
Edema	—	()
Eritema	—	()
Descamación periungueal	—	()
Erupción	—	()
Linfadenopatía cervical	—	()
TOTAL	—	(100)

Matriz de tabla IV.20.- Características demográficas de los casos de síndrome de Kawasaki con brote en Octubre-Diciembre, 1994.

Característica demográfica	Número	Porcentaje
Edad < 1 año	—	()
1 año	—	()
2 años		
3 años	—	()
4 años	—	()
5 años	—	()
≥ 1 año	—	()
Sexo Masculino		
Femenino	—	()
Raza Blanca	—	()
Negra	—	()
Asiática	—	()
Otra	—	()
TOTAL	—	(100)

Matriz de tabla IV.21.- Características epidemiológicas de casos y controles de síndrome de Kawasaki con brote en Octubre-Diciembre, 1994.

Característica epidemiológica		Casos		Controles	
		Número	Porcentaje	Número	Porcentaje
Enfermedad antecedente	Sí	—	()	—	()
	No	—	()	—	()
		Odds ratio=		, 95% CI = (,)	
		$\chi^2=$ __		valor p= __	
Exposición a champú de alfombra	Sí	—	()	—	()
	No	—	()	—	()
		Odds ratio=		, 95% CI = (,)	
		$\chi^2=$ __		valor p= __	

Fuente 19, 20 y 21: Centro para el Control de Enfermedades. Informe de toxi-infecciones alimenticias. Atlanta: CCE, 1989.

La matriz de Tabla IV.19 se centra en seis rasgos clínicos: fiebre, inyección conjuntival bilateral, cambios de la boca, cambios de las extremidades periféricas, erupción cutánea y linfadenopatía cervical. En cada caso se indica el número y el porcentaje.

La matriz de Tabla IV.20 se refiere a las características demográficas: edad, sexo y raza. También se indica el número y el porcentaje.

La matriz de Tabla IV.21 se centra en las características epidemiológicas: antecedentes de enfermedad y exposición a *champú de alfombra*, tanto en casos como en controles. Aquí también se indica el número y el porcentaje.

Los siguientes formatos de matriz de tabla han mostrado ser muy útiles en la investigación de brotes de intoxicación e infecciones alimenticias.

Matriz de tabla IV.22- Encuesta de consumo-enfermedad.

Nombre	Sexo	Edad	Incubación	Síntomas - Alimentos		
				D	V	Obs.

Fuente: Armijo Rojas R. Epidemiología básica en atención primaria de la salud. Madrid: Ediciones Díaz de Santos, 1994.

La matriz de Tabla IV.22 corresponde al formulario “Encuesta de consumo-enfermedad” que se debe llenar para la historia de cada caso de toxi-infección alimenticia. Aquí se incluye nombre, sexo, edad, incubación y síntomas.

Matriz de tabla IV.23- Tabla de consumo-enfermedad.

Alimento	(1) Nº que consumió	(2) Nº que enfermó	(3) Tasa (%)	(4) Nº que no consumió	(5) Nº que enfermó	(6) Tasa (%)	(7) Total comensales

Fuente: Armijo Rojas R. Epidemiología básica en atención primaria de la salud. Madrid: Ediciones Díaz de Santos, 1994.

La matriz de Tabla IV.23 constituye la “Tabla de consumo-enfermedad” del brote epidemiológico de toxi-infección alimenticia. Aquí se especifica, para cada alimento, los que lo consumieron y enfermaron, los que no lo consumieron y enfermaron, con sus tasas (%) respectivas.

8. Gráficos

La Sección “3.2. Gráficos” del Capítulo A “Análisis descriptivo” de la Parte II de este libro (páginas 136-147) introduce el tema de los gráficos señalando sus objetivos y normas de construcción; luego los agrupa en a) gráficos para datos unidimensionales; b) gráficos para datos bidimensionales; c) gráficos para datos multidimensionales.

Introducción

Un gráfico es una manera de mostrar visualmente datos cuantitativos usando un sistema de coordenadas. Es una clase de instantánea estadística que nos ayuda a ver y analizar patrones, tendencias, semejanzas y diferencias en los datos. Además un gráfico es una manera ideal de presentar datos a otras personas. La audiencia de una presentación oral recordará los aspectos importantes de los datos mejor desde un gráfico que desde una tabla.

Tanto en clínica como en epidemiología usamos comunmente gráficos de coordenadas rectangulares, que tienen dos líneas: una horizontal y una vertical las cuales se intersecan en un ángulo recto. Nos referimos a esas líneas como al eje horizontal (eje X) y al eje vertical (eje Y).

Habitualmente usamos el eje horizontal (eje X) para mostrar los valores de la variable independiente la cual es el método de clasificación, tal como el tiempo. Usamos el eje vertical (eje Y) para mostrar la variable dependiente, la cual en epidemiología es usualmente una medida de frecuencia, tal como el número de casos o la tasa de la enfermedad. Rotulamos cada eje para mostrar lo que representa (el nombre de la variable y las unidades en las cuales se mide), y marcamos la escala de medición a lo largo de la línea.

Tanto los gráficos como las tablas completan el texto y ayudan a que el lector tenga una visión y una comprensión rápida del mensaje. Los gráficos y las tablas clarifican la información, ponen énfasis en los datos más significativos, establecen relaciones y resumen el material de los hallazgos ²³.

La producción de documentos técnico-científicos es objeto de un cambio continuo debido, en gran parte, al advenimiento de paquetes de programas por ordenadores. Sea que la producción de las ayudas visuales se hagan manualmente o por ordenador, el autor de una comunicación científica oral o escrita debe comprender los principios básicos de la presentación visual de información, tanto cuantitativa como cualitativa. Esto le permitirá seleccionar las ayudas visuales que mejor comuniquen la información a los lectores y de juzgar la pertinencia y la calidad de los productos utilizados en las diferentes publicaciones biomédicas.

Los principios básicos de la presentación visual de información cuantitativa se pueden resumir como sigue:

- a. Mostrar, hacer visible la información.
- b. Inducir al lector a pensar en lo sustancial en vez de lo intrascendente.
- c. Evitar la distorsión de la información.
- d. Presentar muchos números en un espacio pequeño.
- e. Exponer conjuntos grandes de datos en forma coherente.
- f. Estimular el ojo del lector para comparar diferentes datos, solos o en conjuntos.
- g. Revelar información a diferentes niveles de detalle, desde una visión amplia hasta una estructura fina.
- h. Servir un propósito de descripción, exploración, tabulación y comparación, en forma clara, concisa y exacta.
- i. Integrar estrechamente la imagen de las ideas con la descripción textual²⁴.

Diseño y construcción de gráficos

La comunidad científica internacional también ha alcanzado consenso sobre los fundamentos o bases para la presentación de información en forma de gráficos. Esto incluye elementos importantes tales como: numeración, título, identificación de partes, tamaño y forma, preparación, créditos y permisos, ubicación con respecto al texto²⁵.

1. Numeración

Cuando el documento contenga varios tipos de gráficos (en barra, lineales, en sector) y otro tipo de ilustraciones (diagramas, mapas, radiografías, etc.) es necesario rotularlos todos ellos como «Figuras» y numerarlas consecutivamente usando números árabes.

2. Título

Todo gráfico debe tener un título el cual se presenta en la forma de una frase breve explicativa. Valen aquí los comentarios sobre «Título» de la sección sobre «Tablas».

3. Identificación de partes

Con frecuencia se utilizan en el texto palabras tales como «superior», «inferior», «izquierda», «derecha», «encima», «debajo», «de izquierda a derecha», «en el sentido de los punteros del reloj», y otras más, para ubicar aspectos específicos en un grá-

fico o en partes de un conjunto de gráficos; estas palabras deben ir subrayadas, en itálica o en negrita y deben preceder la frase que identifica el objeto o la persona.

4. Tamaño y forma

El tamaño de un gráfico no debe exceder —como regla general— el de una página de texto. De hecho la mayoría de las veces es posible acomodar dos o más figuras en una página. Los gráficos cuyo tamaño original excede una página pueden habitualmente reducirse fotográficamente. Si la reducción no fuera posible, como en el caso de grandes mapas, el material se puede ubicar en dos páginas cara a cara; otra solución es el insertarla entre dos páginas, efectuando los dobleces necesarios. Valen aquí las observaciones mencionadas en «Tamaño y Forma» del Capítulo «Tablas».

La forma de los gráficos varía según su tipo y es objeto de descripción detallada y de discusión más adelante.

5. Preparación

El arte gráfico ha logrado gran avance creativo y tecnológico, llegando a constituir una especialidad en diversas universidades y empresas comerciales. Cuando exista el recurso especializado y el autor pueda financiarlo, se sugiere utilizarlo sin reserva. La realidad de la vida es, sin embargo, que en la mayoría de los casos el autor deberá preparar sus propios gráficos actuando como aficionado artista y como técnico. Las observaciones que siguen intentan minimizar las consecuencias de tal ventura.

Una consideración preliminar sobre la cual nunca se insistirá demasiado es la necesidad de que el autor tenga absolutamente claro en su mente el objetivo y los principios básicos de cada gráfico; ello le permitirá elegir el tipo más adecuado.

En cuanto a los mapas, éstos suelen estar disponibles en distintos tipos y algunos pueden servir en forma satisfactoria sin mayores adiciones excepto naturalmente, el número de página, número de figura, título o leyenda. Algunos mapas pueden utilizarse como mapas de base, destacando áreas específicas.

Con respecto a las fotografías, se recomienda que sean nítidas, de muy buena calidad; el tamaño 10x15 cm (tamaño de una tarjeta postal) en superficie mate.

El gráfico debe centrarse en la página, con un margen más ancho abajo que arriba para acomodar la leyenda; también el margen izquierdo debe ser mayor que el derecho por razones de empaste.

6. Créditos y permisos

Siempre que se presente material de otras publicaciones, es necesario incluir una frase breve sobre la «Fuente» de la ilustración. Aquí se debe indicar los permisos formales correspondientes de reproducción si la fuente original está protegida por derechos de autor. Si el material utilizado es de derecho público, el permiso no procede; sin embargo subsiste la obligación moral de mencionar la fuente original. En el caso que se utilice material encargado o mandado a hacer por el autor, especialmente mapas, fotografías y dibujos, la cortesía profesional dicta el mencionar al creador del material; esta mención se realiza en la introducción del documento o debajo de cada ilustración.

7. Ubicación

Los gráficos y los esquemas deben ubicarse lo más cerca posible de su primera referencia en el texto. Al igual que las tablas, deben referirse por número de manera que su ubicación exacta sea fácil. En algunos documentos puede haber razones para agrupar las ilustraciones —si son de un mismo tipo— en un apéndice.

Cuando el material ilustrativo es demasiado grande para incluirlo en el texto, éste se ubica en una página entera; la cual se numera consecutivamente con las páginas del texto. No es aceptable el asignar números de página suplementarios (ej: “pág. 17a”). Los mapas se deben numerar en el centro de la parte más alta.

Si aparecen dos o más gráficos en una misma página, cada uno debe llevar un número y una leyenda. También se pueden acomodar dos o más gráficos directamente relacionadas en la misma página como un grupo o conjunto con un único número y leyenda única; en este caso los gráficos individuales se deben identificar con una letra sola (a, b, c, etc.).

Un gráfico ancho se puede acomodar de lado en el página, con la parte superior próxima al empaste. En este caso tanto el número como el título aparecerán debajo del gráfico. El número de la página ocupará su posición normal. Algunas ilustraciones —mapas en particular— llevan impreso dentro de ellas un título o encabezamiento. En tal caso se debe colocar la palabra «Figura» y el número correspondiente de la ilustración en el centro, debajo de la ilustración. Si se incluye una clave o escala (kilómetros, metros, centímetros, milímetros, gramos, miligramos, etc.), ésta debe colocarse convenientemente en un espacio al lado o dentro de la ilustración.

Gráficos en barra

Los gráficos en barra, también denominados gráficos en columna, son las ayudas visuales de elección para comparar magnitudes o tamaños; también son excelentes para destacar diferencias en un *item* en diferentes períodos de tiempo.

1. Uso

Los gráficos en barra muestran específicamente: a) mediciones de una variable o característica en diferentes tiempos o intervalos de tiempo; b) mediciones de diferentes características o atributos durante el mismo tiempo; c) mediciones de diferentes partes de una variable, la cual constituye un todo.

Se recomienda:

- a. Utilizar gráficos en barra simple para explicar relaciones simples.
- b. Utilizar gráficos en barra subdividida (segmentados) para comparar totales y sumas de totales a lo largo de un período de tiempo.
- c. Utilizar gráficos en barra de componente 100% para comparar los componentes de un total en un momento.
- d. Utilizar gráficos en barras agrupadas para comparar varios aspectos de diferentes *items* en un momento.
- e. Utilizar gráficos pareados para comparar dos tipos diferentes de datos.
- f. Utilizar gráficos de desviación para mostrar valores positivos y negativos en un momento.
- g. Utilizar pictográficos, los cuales son una variación del gráfico en barra, para mostrar el elemento que se presenta o se compara. Las barras se hacen con figuras simbólicas de hombres y mujeres, de niños, de camas de hospital, de bebé en una balanza, de píldoras, de jeringas, etc., según el caso.

La utilización de pictográficos ha ganado más popularidad en periódicos, semanarios, afiches, etc. que en publicaciones científicas.

2. Diseño

El diseño de un gráfico en barra debe ser simple, lógico y consistente. A efectos de una fácil lectura, se recomienda que no tenga más de 10 barras. Las variables y las constantes deben rotularse con el lenguaje más directo y más simple posible. La cantidad de datos a acomodar en una página y el tamaño recortado de la página afecta el trazado del gráfico en barras.

La información que sigue ha probado ser útil para la decisión sobre el diseño de un gráfico en barras:

- a. Se debe decidir si las barras serán verticales u horizontales. Hay autores que recomiendan barras verticales para mostrar altitud, mientras que distancias recorridas se muestran mejor con barras horizontales; un arreglo vertical de las barras ilustra mejor la cantidad de una variable en diferentes intervalos de tiempo mientras que un arreglo horizontal de las barras ilustra mejor diferentes mediciones de varias variables en el mismo tiempo.
- b. La escala de medición de las barras debe ser consistente.
- c. Las barras deben tener el mismo ancho. Se sugiere un ancho entre 0,5 y 1 centímetro, aun cuando la altura y el espacio disponible suelen determinar el ancho de las barras.
- d. Las barras deben arreglarse en orden descendiente, con la barra más larga a la izquierda y la más corta a la derecha. Si el tipo de información a mostrar no hace posible este arreglo, hay que asegurarse que el lector comprende la lógica del arreglo que se presenta.
- e. Las barras se deben separar manteniendo igual distancia entre ellas. El autor debe sentirse en libertad para agrupar un conjunto de barras que representan variables con algún elemento en común; si éste fuera el caso se debe mantener igual distancia entre conjuntos de barras.
- f. Se debe poner rótulos en el lado izquierdo o debajo de las barras, aparte del eje del gráfico. Nunca poner rótulos dentro de las barras.
- g. En caso de utilizar color en las barras, ubicar los valores más oscuros a la izquierda y los más pálidos a la derecha. Si se usa un solo color, asegurarse que los tonos entre barras cambian por lo menos en un 20% a fin de mantener una diferencia entre las barras para lograr una impresión clara. Esto limita el número de barras a cinco.
- h. Si un gráfico en barras carece de proporciones, los datos pueden aparecer distorsionados (minimizados o exagerados). Por consiguiente hay que asegurarse que las barras son más anchas que los espacios entre ellas y que las barras tienen la longitud y el grosor adecuado; además hay que recordar que las barras altas y delgadas tienden a exagerar las unidades de medida representadas mientras que las barras bajas y gruesas tienden a minimizar las unidades de medida.

3. Ejemplos de gráficos en barra

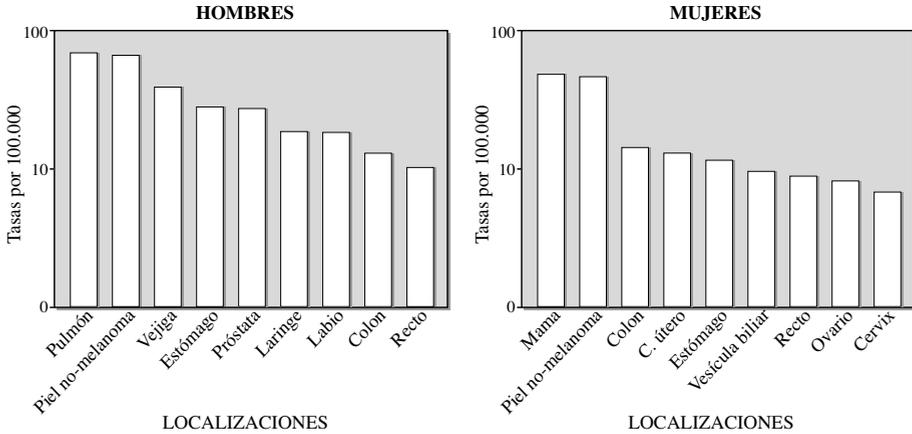


Gráfico IV.1. Incidencia de Cáncer. Localizaciones más frecuentes por sexo. Tasas brutas. Granada 1988-90.

Fuente: Martínez García, C. El Cáncer en Granada, Incidencia y mortalidad 1988-90. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.1 de barras verticales, muestra la incidencia de cáncer en Granada según localizaciones más frecuentes, por sexo.

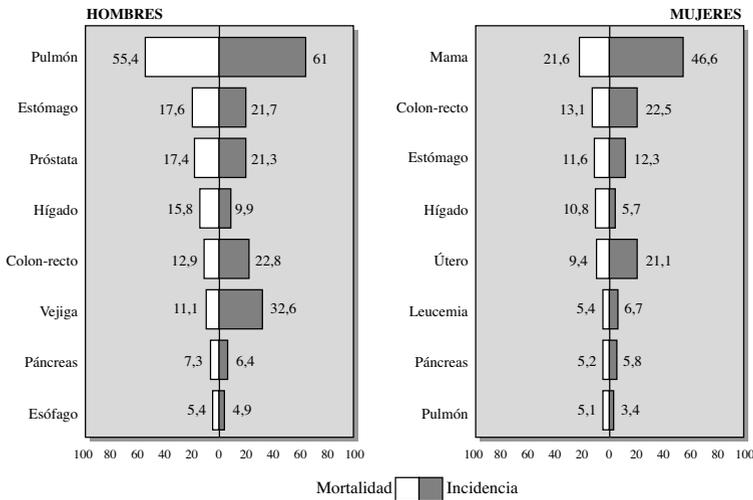


Gráfico IV.2. Mortalidad e incidencia de Cáncer. Tasas brutas por 100.000. Granada 1988-90.

Fuente: Martínez García, C. El Cáncer en Granada, Incidencia y mortalidad 1988-90. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.2 de barras horizontales muestra mortalidad e incidencia de cáncer en Granada según localizaciones más frecuentes, por sexo.

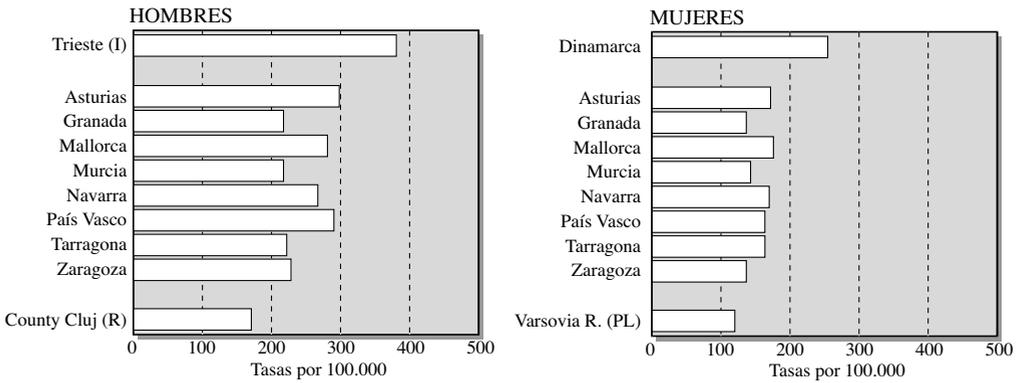


Gráfico IV.3. Incidencia de cáncer excepto Piel-no Melanoma. Comparación de los registros españoles con los valores máximo y mínimo europeos. Tasas estandarizadas por la población mundial. 1988-90.

Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90.* Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.3 de barras horizontales muestra la incidencia de cáncer, excepto piel no melanoma, por sexo.

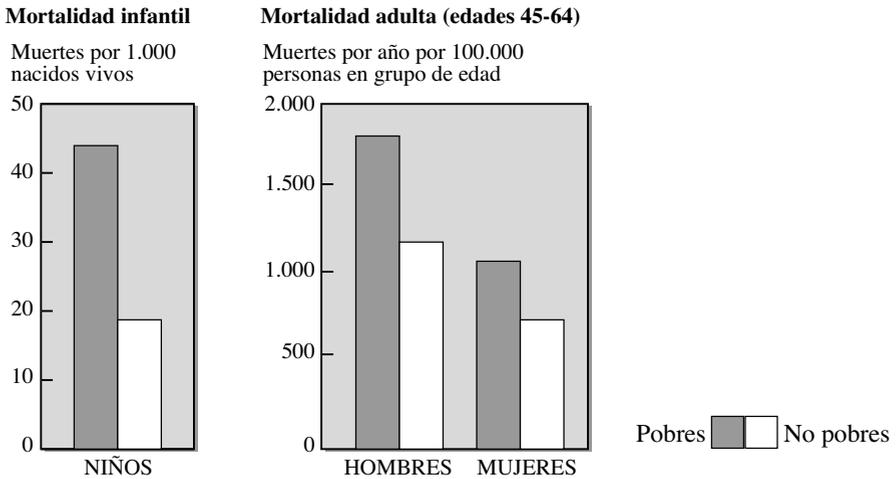


Gráfico IV.4. Mortalidad infantil y adulta en barrios pobres y no pobres de Porto Alegre, Brasil 1990.

Fuente: World Bank. *World development report 1993. Investing in health.* New York: Oxford University Press, 1993.

El Gráfico IV.4 de barras verticales agrupadas muestra mortalidad infantil y mortalidad adulta en barrios pobres y barrios afluentes.

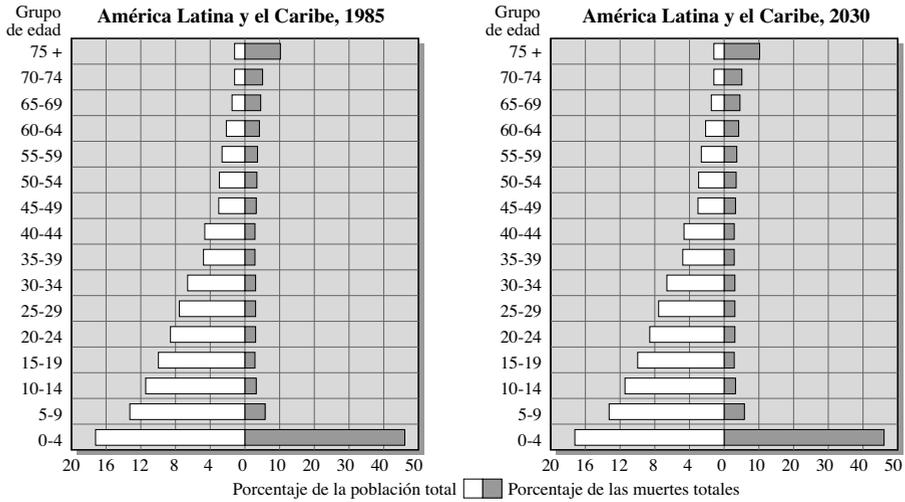


Gráfico IV.5. Patrón cambiante de distribución de edad y mortalidad en América Latina y el Caribe. 1985 y 2030.

Fuente: World Bank. *World development report 1993. Investing in health*. New York: Oxford University Press, 1993.

El Gráfico IV.5 en barras horizontales pareadas muestra el patrón evolucionante de distribución de edad y de mortalidad en América Latina y el Caribe.

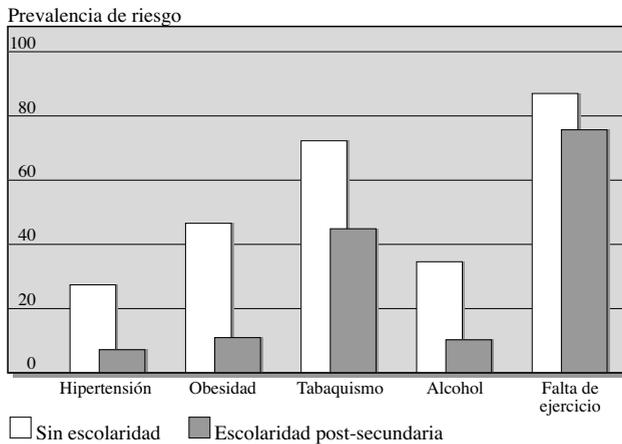


Gráfico IV.6. Escolaridad y factores de riesgo para la salud del adulto en Porto Alegre, Brasil. 1987.

Fuente: World Bank. *World development report 1993. Investing in health*. New York: Oxford University Press, 1993.

El Gráfico IV.6 en barras verticales agrupadas y montadas de a dos muestra la relación entre escolaridad y cinco factores de riesgo en adultos en Porto Alegre, Brasil.

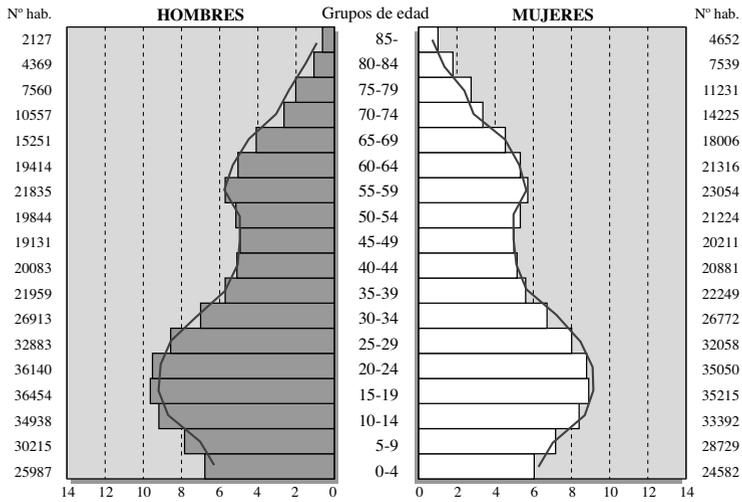


Gráfico IV.7. Pirámide de población de Granada (Media 1988-90)
(En líneas se representa la población española, Censo 1991).

Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

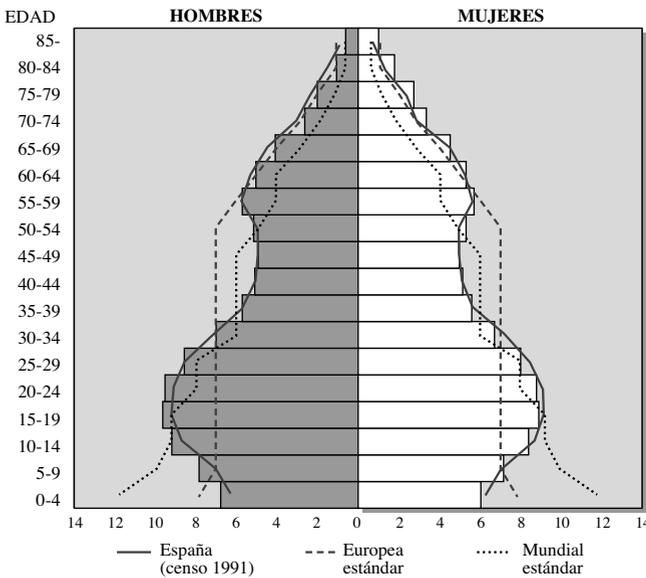


Gráfico IV.8. Pirámide de población de Granada (Media 1988-90)
Comparación con diversas poblaciones.

Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.7 en barras horizontales pareadas separadas muestra la pirámide de población de Granada, por sexo. El IV.8 añade la comparación con España, Europa estándar y mundial estándar.

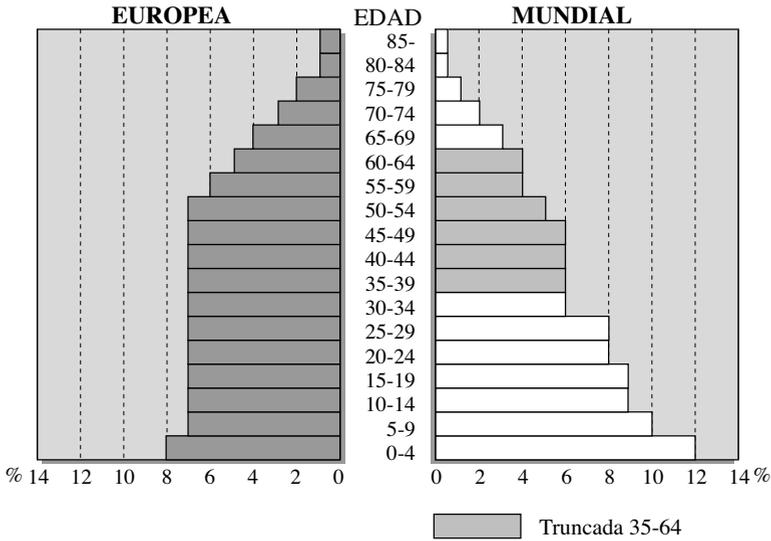


Gráfico IV.9. Poblaciones estándar.

Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

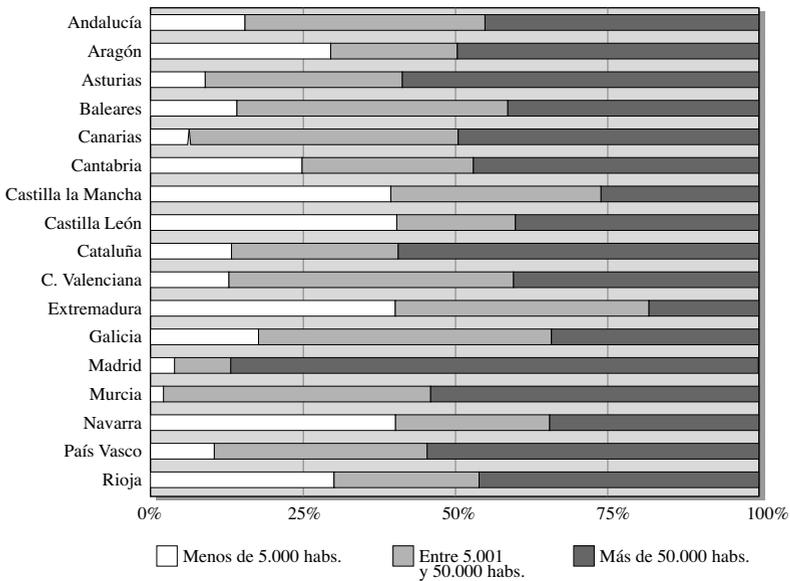


Gráfico IV.10. Población de derecho según tamaño de municipios por comunidades autónomas. España, 1991.

Fuente: Ministerio de Sanidad y Consumo. *Estudio territorializado de la evolución del sector salud a partir de 1980*. Madrid: Ministerio de Sanidad y Consumo, 1995.

El Gráfico IV.9 en barras horizontales pareadas separadas muestra las poblaciones estándar europea y mundial.

El Gráfico IV.10 en barras horizontales de componente 100%, muestra los componentes de la población de derecho según tamaño de los municipios de las 17 comunidades autónomas de España.

Histogramas

Un histograma es un gráfico de distribución de frecuencias de una variable continua. El histograma es el método más tradicional de representar una distribución de frecuencias. Utiliza columnas contiguas para representar el número de observaciones para cada intervalo de clase en la distribución. El área de cada columna es proporcional al número de observaciones en ese intervalo. El área de un histograma es igual al producto de la base del rectángulo (que recoge la amplitud del intervalo de clase) por la altura (que recoge la frecuencia).

1. Uso

Se recomienda utilizar intervalos de clase iguales; de esta manera la altura de la columna está en proporción al número de observaciones que muestra. No se recomienda utilizar histogramas con intervalos de clase desiguales por ser difíciles de construir y de interpretar.

Tampoco se recomienda utilizar escalas quebradas en el eje vertical (eje Y) del histograma porque dan un cuadro engañoso de las frecuencias relativas.

La variable del eje horizontal (eje X) más común es tiempo. Sin embargo, otras variables continuas, tales como nivel de colesterol, nivel de presión arterial, valores de seroalbúmina en gramos por litro de sangre, casos de hepatitis A reportados, etc., pueden inscribirse en el eje X.

Los epidemiólogos frecuentemente crean y discuten *curvas epidémicas*. Una curva epidémica no es una curva del todo sino un histograma el cual muestra los casos de enfermedad durante el brote o la epidemia desde la fecha de comienzo. Frecuentemente se diseñan las columnas como pilas de cuadrados, donde cada cuadrado representa un caso. En una curva epidémica cada número debería centrarse entre las marcas señales del intervalo apropiado.

Se utiliza cualquier intervalo de tiempo que sea apropiado para la enfermedad en cuestión: tal vez horas para un brote de gastroenteritis por *C. perfringens*; tres a cinco días para un brote de hepatitis A. Como regla general se recomienda hacer los intervalos menores de un cuarto del período de incubación de la enfermedad a mostrar.

Se recomienda empezar el eje X antes del primer caso del brote y mostrar cualquier caso de la misma enfermedad que ocurrió durante el período preepidémico. Esos casos pueden representar casos antecedentes o casos no relacionados. Al mismo tiempo pueden representar la fuente del brote.

2. Diseño

En la construcción de un histograma se debe tener presente los siguientes elementos:

- a. El eje de la abscisa (horizontal) es una escala continua donde se reflejan todas las unidades que agrupan los intervalos de clase.
- b. Para cada clase en la distribución se dibuja un rectángulo vertical extendiéndose desde el valor inferior del intervalo hasta el superior.
- c. Las áreas de estos rectángulos deben ser proporcionales a la frecuencia de la clase.
- d. No pueden existir separaciones entre los rectángulos, dado que, al ser continua la variable, los límites de los intervalos de clase son límites que reflejan valores reales.
- e. Cuando los intervalos son desiguales (unos más amplios que otros), debe respetarse la proporcionalidad entre las áreas, lo que afecta la comparabilidad entre barras; ya hemos considerado cómo bajo estas circunstancias la comparación entre áreas se hace más dificultosa.
- f. Cuando los intervalos son iguales, las bases de los rectángulos también lo son, por lo que la atención se centra en la altura casi exclusivamente, a efectos de comparación, dado que la altura del rectángulo será proporcional a la frecuencia.

3. Ejemplos de histogramas

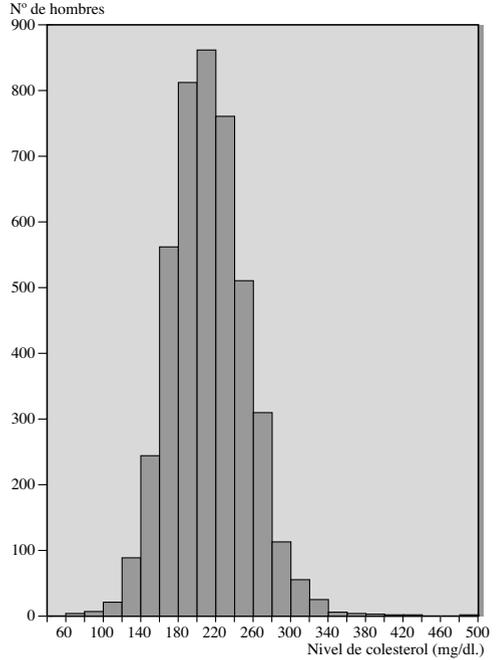
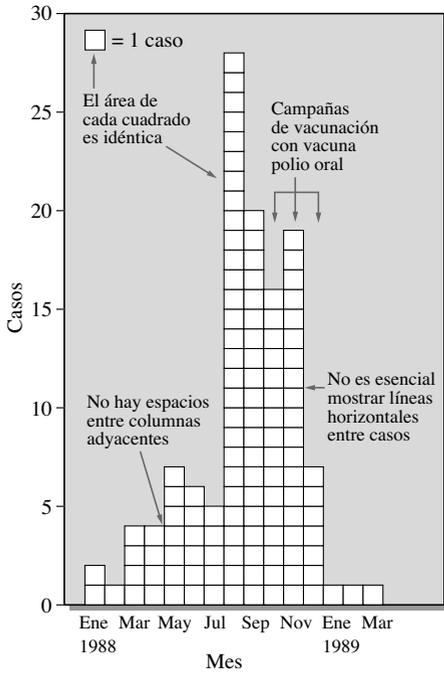


Gráfico IV.11. Ejemplo de histograma. Casos reportados de poliomielitis parálitica según mes de ocurrencia. Enero 1988-Marzo 1989.

Fuente: Centro para el control de enfermedades. Informe de Toxi-infecciones alimenticias. Atlanta: CCE, 1989.

Gráfico IV.12. Ejemplo de histograma. Niveles de colesterol reportados entre 4.462 hombres. Estados Unidos, 1985-1986.

Fuente: Centro para el control de enfermedades. Informe de Toxi-infecciones alimenticias. Atlanta: CCE, 1989.

El Gráfico IV.11 es un histograma sobre casos reportados de poliomielitis parálitica según mes de ocurrencia. Se recalca algunos aspectos de diseño de histogramas, tales como tamaño de los cuadrados, falta de espacio entre columnas adyacentes y líneas horizontales entre casos.

El Gráfico IV.12 es un histograma que muestra niveles de colesterol reportados en un estudio de salud en hombres en Estados Unidos. El eje Y indica el número de hombres con escala de intervalo de cien. El eje X muestra los niveles de colesterol (mg/dl) con intervalos de escala de 20 mg/dl.

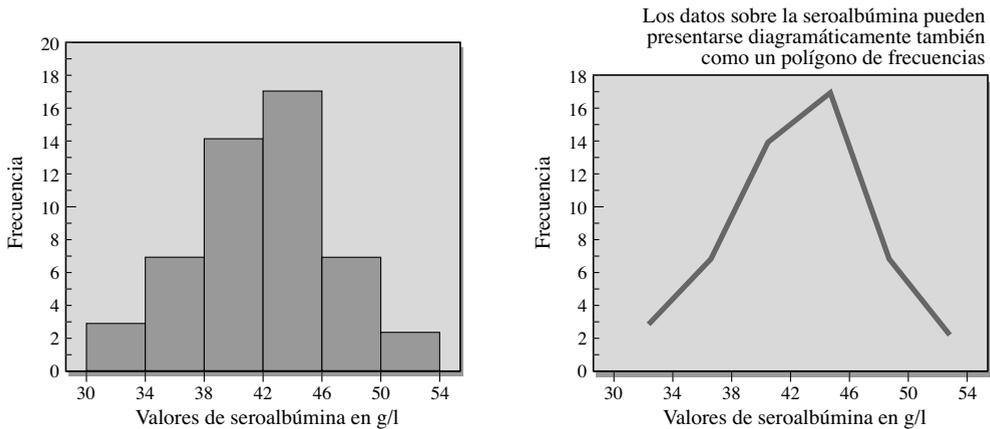


Gráfico IV.13. Histograma y polígono de frecuencias de los datos sobre seroalbúmina.

Fuente: Lwanga SK, Tye CY. La enseñanza de la estadística sanitaria. Ginebra, 1987.

El Gráfico IV.13 es un histograma de datos de valores de seroalbúmina, construido sobre la información de la tabla IV.3 (ver apéndice III, Tabla IV.3). La mitad inferior de la página muestra la misma información mediante un polígono de frecuencias.

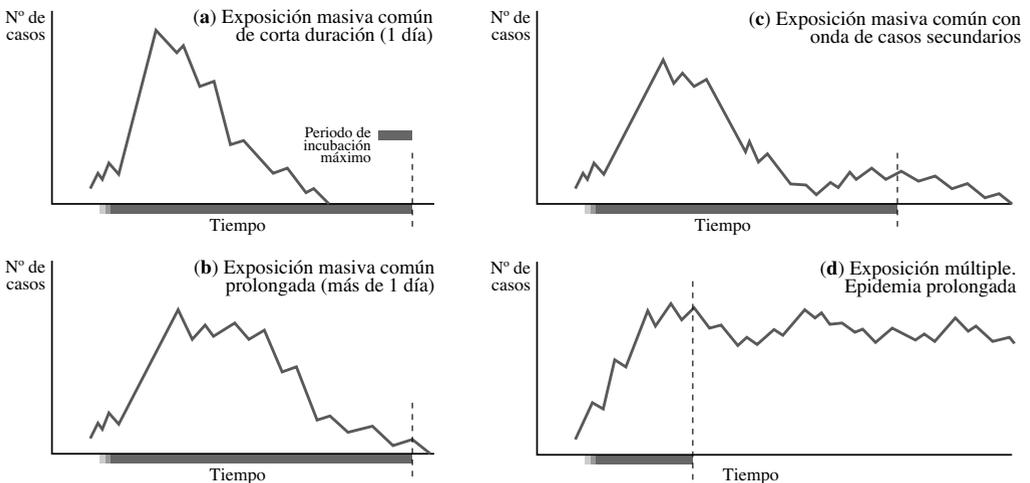


Gráfico IV.14. Representación esquemática de curvas epidémicas.

Fuente: Armijo Rojas R. Epidemiología básica en atención primaria de salud. Madrid: Ediciones Díaz de Santos, 1994.

El Gráfico IV.14 es una representación esquemática de cuatro tipos de curvas epidémicas. El eje Y muestra el número de casos; el eje X muestra el tiempo.

Polígonos de frecuencias

Un polígono de frecuencias, al igual que un histograma, es un gráfico de distribución de frecuencias. En un polígono de frecuencias se marcan las observaciones dentro de un intervalo con un solo punto ubicado en el punto céntrico del intervalo; luego se conecta el conjunto de puntos con una línea recta. El efecto final de unir esta serie de líneas rectas es el de un *polígono*: figura limitada por varias líneas rectas.

1. Uso

Los polígonos de frecuencias facilitan el representar y el comparar dos o más distribuciones sobre el mismo conjunto de ejes.

Un polígono de frecuencia se diferencia de un gráfico lineal de escala aritmética. En efecto: a) usamos polígonos de frecuencia (o histogramas) para representar la distribución de frecuencias enteras de una variable continua; usamos un gráfico lineal de escala aritmética para trazar una serie de puntos de datos observados (cuentas o tasas), habitualmente sobre tiempo; b) un polígono de frecuencia debe cerrarse en ambos extremos debido a que el área bajo la curva es representativa de los datos; un gráfico lineal de escala aritmética simplemente une los puntos de los datos mediante una línea.

Así como el histograma refleja una distribución de frecuencias, el polígono de frecuencias facilita la comparación entre dos o más distribuciones de frecuencias.

2. Diseño

- a. La línea de la frecuencia del polígono se mueve desde un punto céntrico a otro punto céntrico, creando series de pares de triángulos de igual tamaño: uno situado fuera del histograma y otro situado dentro del histograma. Este es un aspecto necesario de los polígonos de frecuencias: un polígono de frecuencias de un conjunto de datos debe incluir la misma área del histograma de esos datos; en consecuencia por cada área del histograma que el polígono de frecuencia deja afuera, debe “importar” un área de igual tamaño.
- b. Para mantener un área total igual se debe poner especial atención a como se “cierra” el polígono de frecuencias. La línea de frecuencia del polígono empieza en el punto céntrico del intervalo anterior al intervalo uno (con un valor Y de cero), completamente fuera del histograma y conecta con el punto céntrico del primer intervalo que contiene observaciones. El lado derecho del polígono de frecuencia cierra de manera similar.

3. Ejemplos de polígonos de frecuencias

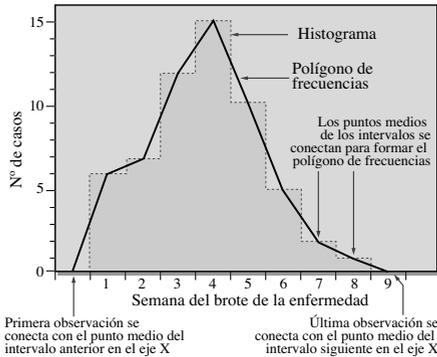


Gráfico IV.15. Número de casos de gripe reportados por semana del brote.

Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

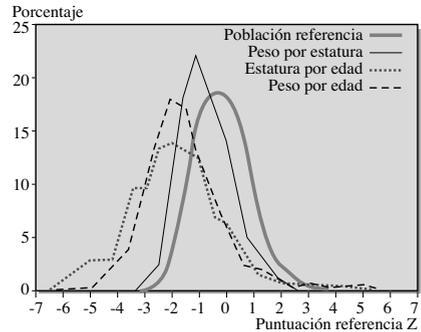


Gráfico IV.16. Antropometría de niños haitianos de edad 24-60 meses comparado con población de referencia OMS.

Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

El Gráfico IV.15 es un polígono de frecuencias construido sobre el perfil de un histograma, para los mismos datos; corresponde a la semana de comienzo de una influenza.

El Gráfico IV.16 es un polígono de frecuencias en el cual se comparan tres polígonos de frecuencia entre sí y contra una curva de distribución normal.

Gráficos lineales de escala aritmética

El gráfico lineal de coordenadas rectilíneas de escala aritmética es el más antiguo, el más simple, el más familiar y el más ampliamente utilizado tanto en clínica como en salud pública. También es el tipo de gráfico para el cual se han formulado los principios de diseño y las estándares más rigurosos. Sin embargo, y a pesar de esos hechos, un número elevado de este tipo de gráficos que aparecen en artículos originales publicados muestran superficialidad, error e incompetencia. En efecto es frecuente observar una planificación y diseño pobre, rótulos difíciles de leer, carencia de líneas en la red y de puntos de la escala, ejecución mediocre de las curvas, títulos poco claros e incompletos, etc.

Los gráficos lineales se utilizan para mostrar la relación entre X e Y como un conjunto de mediciones discretas. En un gráfico lineal la constante se traza horizontalmente a lo largo del eje X y las variables se trazan verticalmente a lo largo del eje Y. El gráfico lineal muestra una serie de mediciones exactas y tendencias creadas por esas mediciones; en consecuencia un gráfico lineal es adecuado para presentar datos que deben registrarse como una serie de hechos.

1. Uso

Se usan gráficos lineales con cualquier tipo de audiencia para mostrar una tendencia, más que una comparación. Los gráficos lineales son los gráficos más utilizados debido a que permiten mostrar datos discretos como parte de una tendencia y permiten llamar la atención sobre una predicción. Se usa el gráfico lineal cuando se quiere poner énfasis en datos sin distorsionar su efecto. Un buen gráfico lineal puede reemplazar páginas de explicación, permitiendo al lector la comprensión de ideas complejas en unos pocos minutos.

2. Diseño

Dado el hecho que el gráfico lineal representa una relación de proporción entre X e Y, la representación visual debe sugerir y realzar esa proporción.

La información siguiente constituye una ayuda para diseñar un buen gráfico lineal:

- a. Crear los ejes X e Y en forma lógica. Dibujar el eje horizontal X con intervalos iguales de izquierda a derecha, partiendo con el número más bajo a la izquierda. Dibujar el eje vertical Y con cero en el fondo y con el número más alto en la cima. Recordar que su número más bajo será donde X e Y se intersecan. Se debe mantener una proporción constante en esas dos líneas de manera tal que las líneas den un cuadro lógico de los datos. Si no se mantiene una proporción constante o si el dibujo es demasiado ancho o demasiado alto, los datos aparecerán distorsionados porque las líneas aparecen muy empinadas o muy planas.
- b. Mostrar cero como punto de partida para ambos, X e Y, siempre que sea posible. Si razones de espacio exigen comenzar con un número más alto, hacerlo claramente con un espacio en blanco y un marcador que muestre donde empiezan los intervalos; esta técnica se denomina de *cero suprimido*.
- c. Rotular claramente los intervalos de los ejes X e Y, preferiblemente en forma horizontal. Usar una leyenda para identificar los puntos que se trazan e identificar las líneas que se dibujan para mostrar tendencias. Identificar todos los hechos que se suponen importantes para el lector y colocar su clave de identificación en la esquina inferior derecha del gráfico. De esta manera se reserva la esquina inferior izquierda para información sobre «Fuente» y otras notas a pie de gráfico.
- d. Utilizar líneas claramente contrastantes (líneas gruesas, líneas delgadas, líneas quebradas, líneas de punto) en el caso de que el gráfico presente comparaciones múltiples y requiera —en consecuencia— más de una línea de tendencia. Para gráficos en colores, utilizar colores contrastantes.

- e. Utilizar una «rejilla» o «sombreado» para dar mayor exactitud y para realzar la calidad del gráfico. Sin embargo se debe tener cuidado que este artificio no estorbe y confunda el diseño.
- f. Evitar el insertar líneas múltiples para mostrar relaciones entre muchas variables en un solo gráfico. Demasiadas líneas confunden al lector; la mayoría de los artistas gráficos aconsejan no incluir más de 3 a 4 líneas.

3. Ejemplos de gráficos lineales de escala aritmética

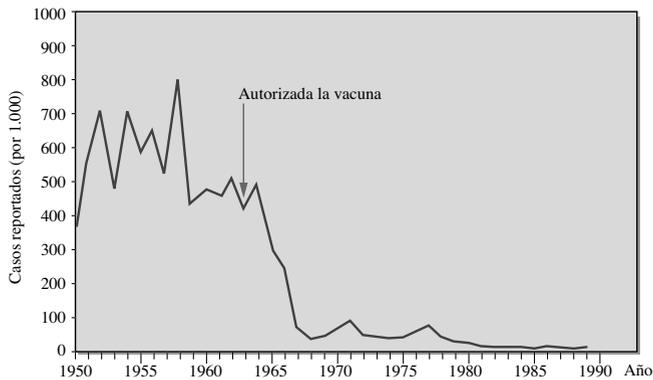


Gráfico IV.17. Casos de sarampión reportados en Estados Unidos. 1900 a 1985.

Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

El Gráfico IV.17 es lineal de escala aritmética muestra la tendencia de casos reportados de sarampión en los Estados Unidos , 1950-1990.

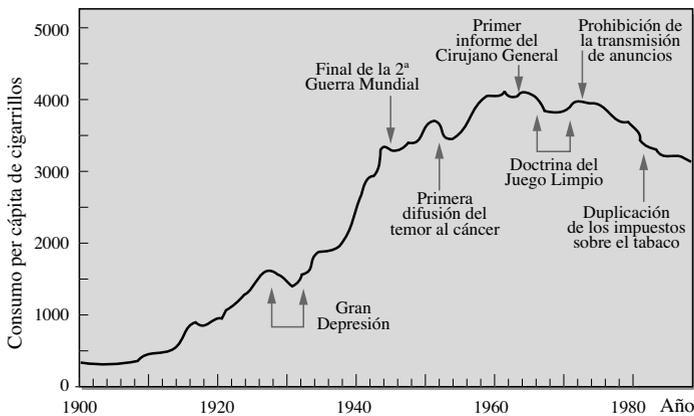


Gráfico IV.18. Consumo per cápita de cigarrillos en adultos y acontecimientos mundiales fundamentales que guardan relación con el hábito de fumar.

Fuente: Murphy G, Lawrence W, Lenhard R. *Oncología clínica. Manual de la American Cancer Society*. 2ª ed. Washington, D.C.: Organización Panamericana de la Salud, 1996. (Tomado de *Reducing the Health Consequences of Smoking*).

El Gráfico IV.18 es lineal de escala aritmética muestra el consumo per cápita de cigarrillos en Estados Unidos 1900 y 1985. Nótese cómo se señala con flechas seis acontecimientos mundiales que se relacionan con el hábito de fumar.

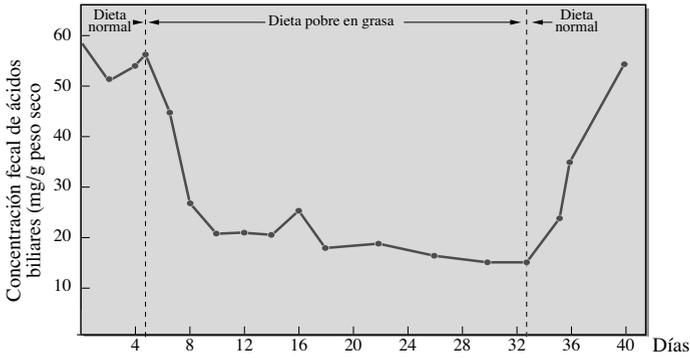


Gráfico IV.19. Efecto de la grasa de la dieta en la concentración fecal de ácidos biliares.

Fuente: Murphy G, Lawrence W, Lenhard R. Oncología clínica. Manual de la American Cancer Society. Segunda ed. Washington, D.C.: Organización Panamericana de la Salud, 1996.

El Gráfico IV.19 es lineal de escala aritmética muestra el efecto de la grasa de la dieta en la concentración fecal de ácidos biliares. Nótese la división del gráfico en tres zonas: antes (dieta normal), durante (dieta pobre en grasa) y después (dieta normal).

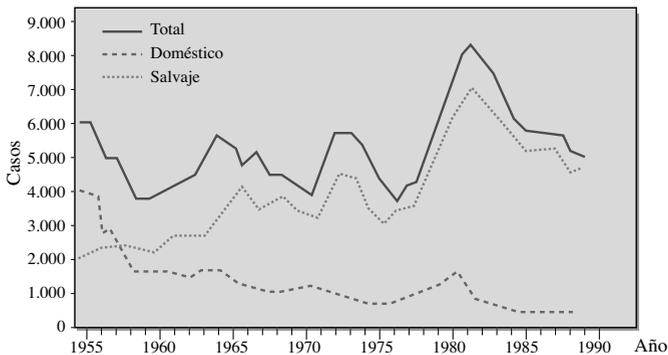


Gráfico IV.20. Casos de rabia doméstica y salvaje según año de reporte. Estados Unidos y Puerto Rico. 1955-1989

Fuente: Centro para el control de enfermedades. Informe de Toxi-infecciones alimenticias. Atlanta: CCE, 1989

El Gráfico IV.20 es lineal de escala aritmética, líneas múltiples, muestra la tendencia de casos de rabia con sus componentes doméstica y selvática en Estados Unidos, 1950-1990.

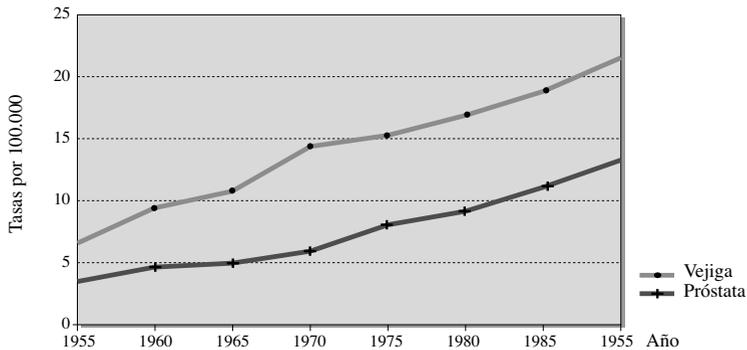


Gráfico IV.21. Tendencias de mortalidad por cáncer de vejiga y próstata. España, 1955-90.

Fuente: Movimiento Natural de la Población. Elaboración: registro de Cáncer de Granada.

El Gráfico IV.21 es lineal de escala aritmética, líneas múltiples, muestra las tendencias de mortalidad, tasas por 100.000 habitantes, por cáncer de vejiga y próstata en España, 1950-1990.

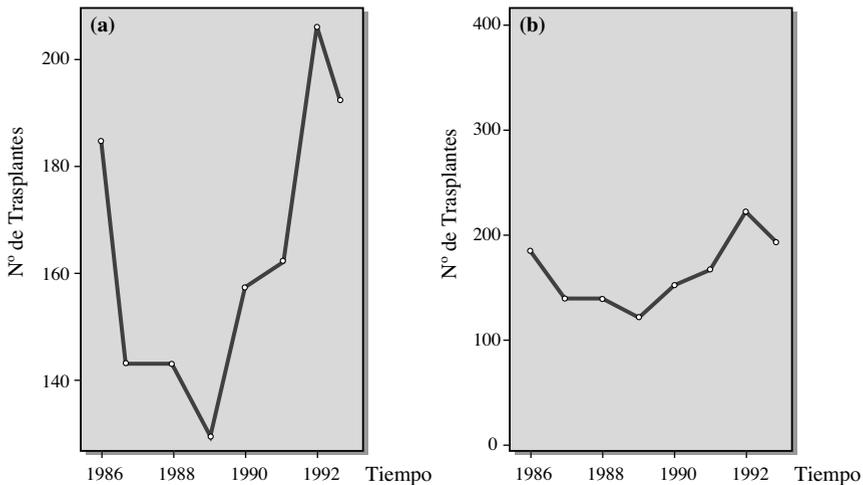


Gráfico IV.22. Trasplantes en Andalucía en el período 1986-93.

La diferencia entre los gráficos (a) y (b) está en la escala del eje vertical.

Fuente: Burgos R, Chicharro JA, Bobenieth MA. *Metodología de Investigación y escritura científica en clínica*. Granada: Escuela Andaluza de Salud Pública, 1996.

El Gráfico IV.22 es lineal de escala aritmética sobre trasplantes en Andalucía 1986-1993. Llama la atención que la diferencia entre (a) y (b) está en la escala del eje vertical (Y), lo cual puede conducir a una distorsión en su interpretación.

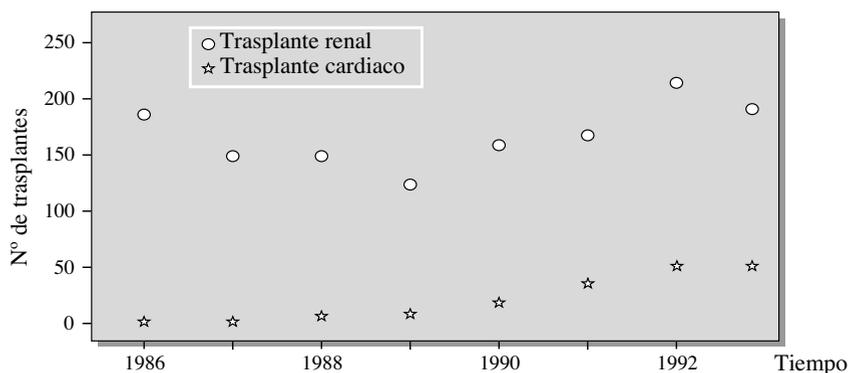


Gráfico IV.23. *Trasplantes renales y cardiacos en Andalucía en el período 1986-93.*

Fuente: Burgos R, Chicharro JA, Bobenrieth MA. *Metodología de Investigación y escritura científica en clínica*. Granada: Escuela Andaluza de Salud Pública, 1996.

El Gráfico IV.23 presenta información sobre trasplantes renales y cardíacos en Andalucía. Tiene la ventaja de presentar las relaciones simultáneamente entre ambas variables, lo que permite observar las fluctuaciones de ambos trasplantes a lo largo del período 1986-1992.

Este gráfico constituye una forma adicional de presentar la información contenida en la Tabla IV.8.

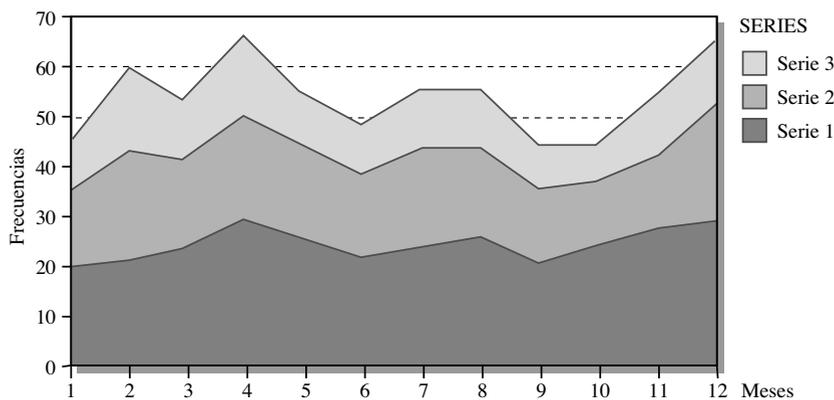


Gráfico IV.24. *Gráfico de estratos.*

Fuente: Alaminos, A. *Gráficos. Cuadernos metodológicos, Nº 7*. Centro de Investigaciones Sociológicas. Madrid: CIS, 1993.

El Gráfico IV.24 es lineal de escala aritmética de líneas compuestas de estratos, en el que los datos comparados son componentes de un agregado de tres series.

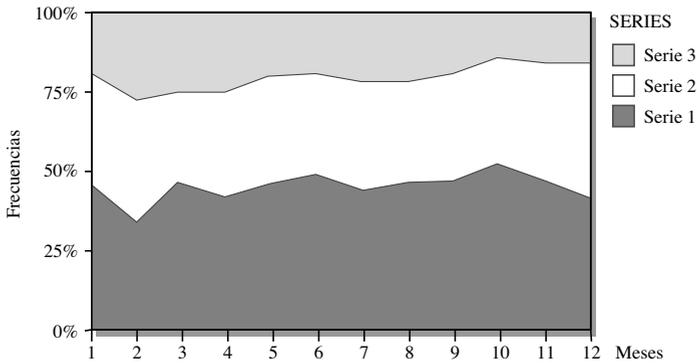


Gráfico IV.25. Gráfico de estratos en base 100%.

Fuente: Alaminos, A. Gráficos. *Cuadernos metodológicos*, Nº 7. Centro de Investigaciones Sociológicas. Madrid: CIS, 1993.

El Gráfico IV.25 es lineal de líneas compuestas o estratos en base 100%, de escala aritmética. La banda rectangular de altura es igual a 100%, donde se representa -según su peso relativa en ese momento- cada una de las tres series.

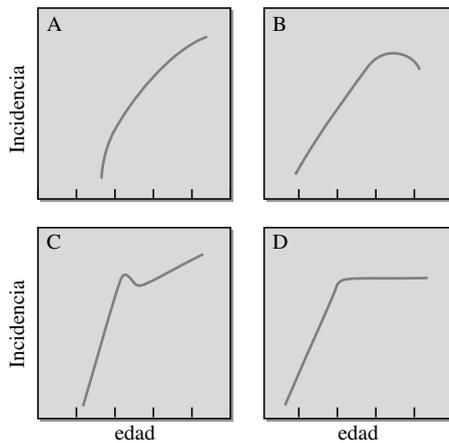


Gráfico IV.26. Patrones de incidencia en cuatro tipos de cánceres cuya evolución se ve afectada por la edad.

Fuente: Martínez Oliva, L. *Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos*. México: Centro Panamericano de Ecología y Salud, 1990.

El Gráfico IV.26 es lineal el cual presenta patrones de incidencia en cuatro tipos de cánceres cuya evolución se ve afectada por la edad.

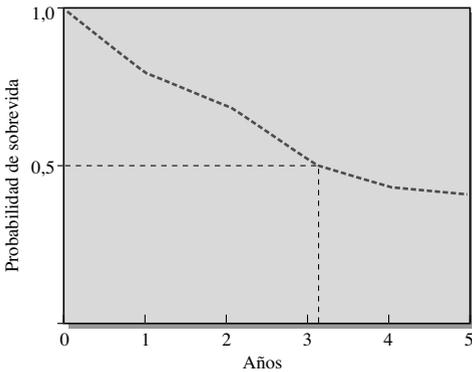


Gráfico IV.27. Curva de sobrevida.

Fuente: Martínez Oliva, L. *Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos.* México: Centro Panamericano de Ecología y Salud, 1990.

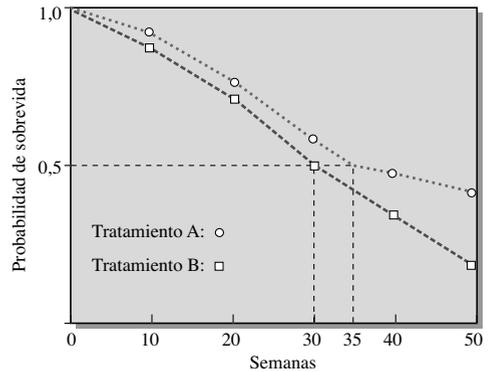


Gráfico IV.28. Curvas de sobrevida hipotéticas.

La sobrevida media es muy similar para ambas curvas, 35 semanas para la curva A y 30 para la curva B; sin embargo, de ahí en adelante ambas experiencias dan resultados muy distintos. A las 50 semanas la probabilidad de sobrevida para A es de 40%, en cambio para B es levemente superior al 10%

El Gráfico IV.27 es lineal y representa una curva de sobrevida. El eje vertical (Y) corresponde a la probabilidad de sobrevida (de 0 hasta 1.0); el eje horizontal (X) corresponde a años (de 1 hasta 5).

El Gráfico IV.28 es lineal y permite comparar dos curvas de sobrevida hipotéticas a las 35 semanas y a las 50 semanas.

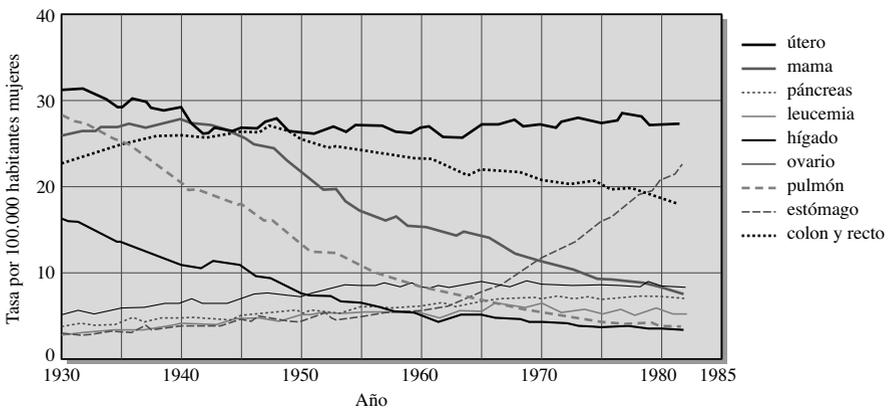


Gráfico IV.29. Tasas de mortalidad por cáncer por 100.000 habitantes para diversas localizaciones, ajustadas para mujeres, por edad, en los Estados Unidos. 1930-1982.

Fuente: Martínez Oliva, L. *Cáncer y ambiente. Bases epidemiológicas para su investigación y control. Métodos epidemiológicos.* México: Centro Panamericano de Ecología y Salud, 1990.

El Gráfico IV.29 es lineal de escala aritmética y presenta tasas de mortalidad por cánceres en nueve localizaciones a lo largo de 55 años, en Estados Unidos. Llama la atención la dificultad para identificar cada localización con el tipo de línea. Esto se debe al error de haber incluido nueve líneas de distinta configuración, pero que al estar todas en blanco y negro inducen a confusión.

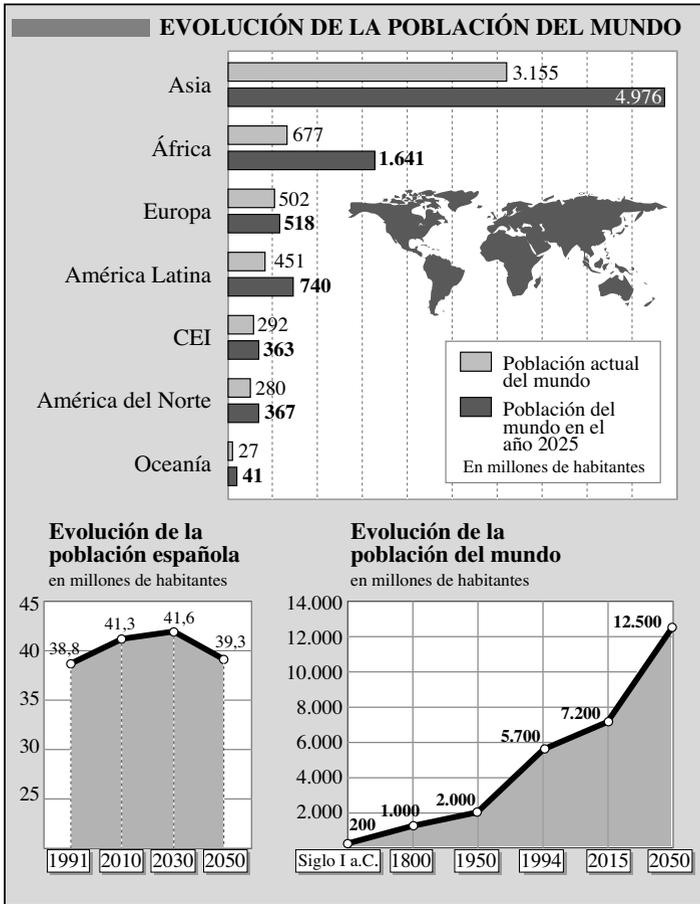


Gráfico IV.30. Evolución de la población del mundo.

Fuente: Naciones Unidas. Estadísticas mundiales. New York: N.U., 1993.

El Gráfico IV.30 presenta la evolución de la población del mundo mediante una combinación de un gráfico de barras horizontales (mitad superior) y un gráfico lineal (cuadrante inferior derecho). Además se muestra un gráfico lineal en silueta sobre la evolución de la población española (cuadrante inferior izquierdo).

Gráficos lineales de escala semilogarítmica

Este tipo de gráfico, también denominados *gráficos de cambio relativo*, son confiables, adaptables y útiles. En efecto permiten comparar cambios relativos para grandes y pequeñas cantidades.

Conviene recordar que en los gráficos lineales de escala aritmética descritos en la sección anterior, espacios iguales representan igual cantidad de *cambio*. En el gráfico semilogarítmico, espacios iguales representan igual cantidad de *cambio relativo*.

Hay cuatro características a tener presente al momento de decidir la utilización de un gráfico semilogarítmico.

- a. El gráfico semilogarítmico es inigualable para comparar tasas proporcionales de cambio para un número de curvas diferentes o para varios segmentos de la misma curva. En contraste el gráfico lineal aritmético compara diferencias o incrementos absolutos de cambio.
- b. El gráfico semilogarítmico pone énfasis en cambio relativo y porcentual. Al mismo tiempo las calibraciones y figuras de escala en el eje vertical provee una referencia rápida de los valores actuales que se han trazado.
- c. El gráfico semilogarítmico representa un medio para comparar con claridad y con confiabilidad dos o más series que se diferencian grandemente en valores absolutos.
- d. El gráfico semilogarítmico provee una técnica sólida para comparar varias series de datos que no se expresan en unidades comunes.

Al parecer el nombre insólito de *semilogarítmico* y la escala logarítmica distintiva de su eje vertical (Y) contribuyen a su condición enigmática. Si a esto se agrega a veces un diseño pobre y una falta de comprensión —aun a nivel superficial— de lo que significa un gráfico semilogarítmico por parte del usuario, el resultado es su falta de utilización en las revistas biomédicas.

1. Uso

Se usan gráficos semilogarítmicos para mostrar cambios relativos cuando las cantidades base difieren grandemente. El gráfico lineal aritmético es útil para describir fluctuaciones comparativas solamente cuando las cantidades comparadas tienen valores similares. También se utilizan gráficos semilogarítmicos para comparar cambios relativos de variables expresadas en unidades diferentes.

2. Diseño

- a. Es fundamental tener presente que el rasgo que comanda el diseño es el énfasis en proporciones o razones de cambio. Por este hecho suele referirse a ellos como a gráficos de proporción. El ejemplo siguiente ayuda a ilustrar este punto. Un aumento de 350 a 400 en un gráfico lineal de escala aritmética mostraría un aumento diez veces mayor en cantidad que un aumento de 35 a 40. En un gráfico lineal de escala semilogarítmica los dos aumentos, medidos por la inclinación de las curvas, serían exactamente iguales debido a que la proporción de sus cambios es idéntica.
- b. La característica estructural más distintiva es su escala o eje vertical Y, el cual es logarítmico.
- c. La escala o eje horizontal X es aritmética, lo cual significa que los intervalos se espacian en forma igual.
- d. El rayado semilogarítmico se compone de uno o más «ciclos», «fases» o «filas». Un ciclo cubre una extensión o radio de acción representado por 10 elevado a la potencia 1 a 10 elevado a la potencia 2 o sea, desde 10 a 100; dos ciclos conforme a la secuencia precedente acomodaría valores desde 10 elevado a la potencia 1 a 10 elevado a la potencia 3 o sea, desde 10 a 1000. Teóricamente, el número de ciclos se puede extender indefinidamente en ambas direcciones y las extensiones se expresan siempre como múltiplos o divisiones de 10.
- e. El número de ciclos, en la práctica, se determina por los valores más altos y más bajos de la serie a trazar. Si la extensión de los valores en una serie es relativamente angosta es adecuado el usar solamente una parte de un ciclo.
- f. No existe línea basal cero en la escala vertical Y; en consecuencia la interpretación de las curvas no está condicionada por la distancia desde la línea base, como es lo típico en un gráfico de escala aritmética.
- g. Como consecuencia del punto anterior, en un gráfico semilogarítmico no es posible pasar de números positivos a números negativos.
- h. Al leer un gráfico semilogarítmico se debe recordar que las distancias verticales indican cambios porcentuales. Por lo tanto el espacio entre 1 y 2 (o entre 10 y 20 o entre 100 y 200) siempre representa un cambio del 100%, independientemente de dónde se mida en el gráfico.

3.Ejemplos de gráficos lineales de escala semilogarítmica

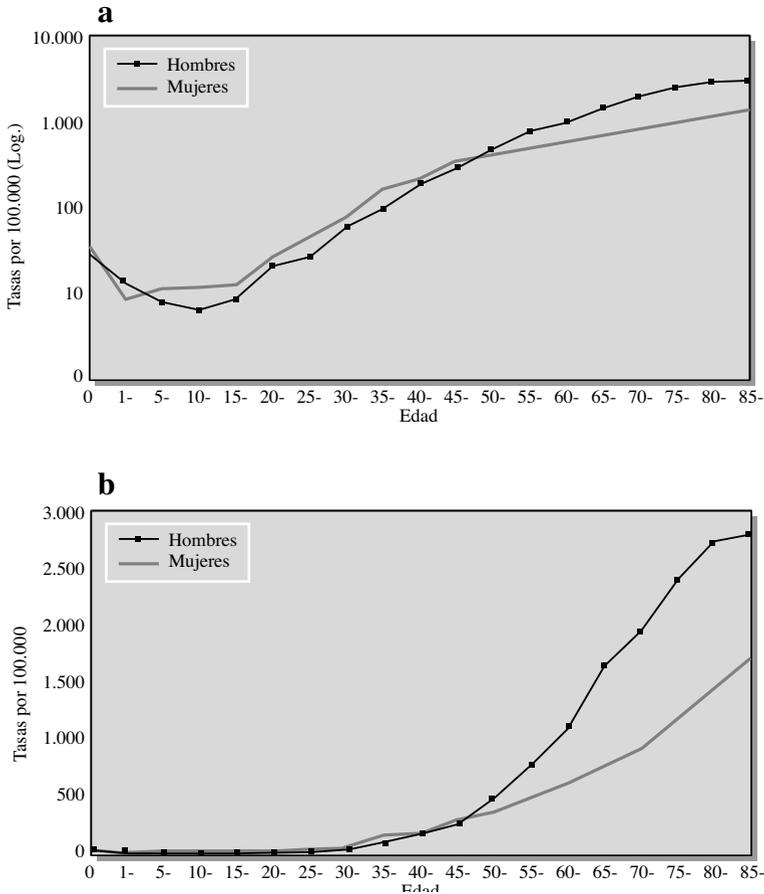


Gráfico IV.31. Incidencia de cáncer todas las localizaciones. Tasas específicas por edad y sexo. Granada 1988-90. Arriba, escala vertical semilogarítmica. Abajo, escala aritmética. Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.31a es lineal de escala semilogarítmica y muestra la incidencia de cáncer en Granada, todas las localizaciones, por sexo.

El gráfico IV.31b muestra la misma información pero en un gráfico lineal de escala aritmética. La diferencia en la escala de configuración de ambos gráficos es notable.

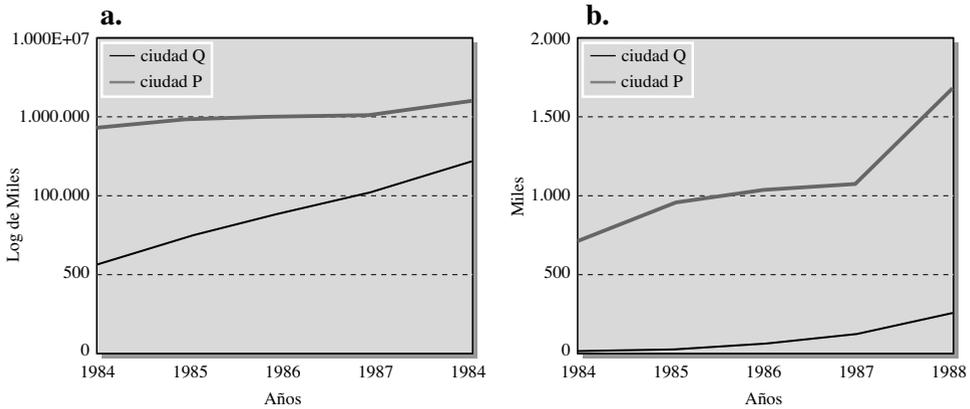


Gráfico IV.32. Evolución de la población en dos ciudades.

Izquierda, escala vertical logarítmica. Derecha, escala aritmética.

Fuente: Alaminos, A. Gráficos. *Cuadernos metodológicos*, N° 7. Centro de Investigaciones Sociológicas. Madrid: CIS, 1993.

El Gráfico IV.32 es lineal y muestra la evolución de la población en dos ciudades. El gráfico IV.32a. tiene el eje Y con escala logarítmica. El gráfico IV.32b. tiene el eje Y con escala aritmética, lo cual produce una distorsión manifiesta.

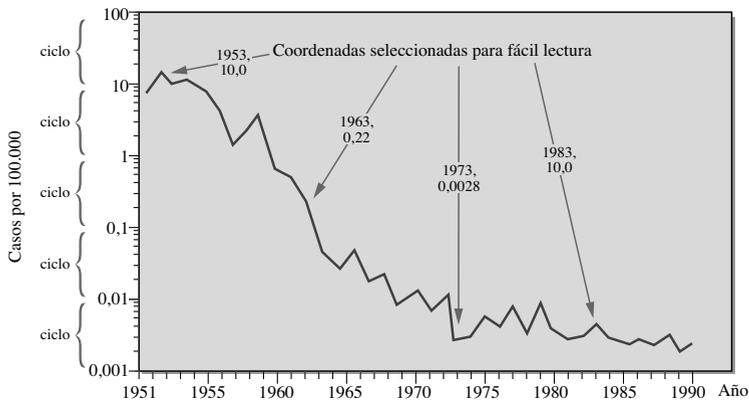


Gráfico IV.33. Casos reportados de poliomielitis parálitica, por 10⁵ habitantes y año. Estados Unidos, 1951-1989.

Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

El Gráfico IV.33 es lineal de escala semilogarítmica y muestra casos de poliomielitis parálitica por 100.000 habitantes. Nótese: a) la escala Y logarítmica se ha dividido en 5 ciclos, cada uno de los cuales cubre igual distancia; b) el trazado de las frecuencias se señala con cuatro coordenadas seleccionadas para una más fácil lectura.

Gráficos en sectores

Los gráficos en sectores, llamados también gráficos *circulares* o *gráficos en tarta*, son los gráficos más simples disponibles al autor de un artículo. El gráfico circular —al igual que una tarta— se divide en «sectores» o «cuñas». Cada sector se rotula para indicar un porcentaje de la cantidad total (100%). El gráfico circular realza el texto y pone énfasis en relaciones importantes entre los hechos presentados.

1. Uso

Su uso está indicado para grupos de audiencias con intereses y antecedentes sobre el tema muy variados. Se los utiliza por su rápida facilidad de lectura y comprensión. Por su naturaleza, este tipo de gráfico da a la audiencia una visión general, un cuadro panorámico sobre valores básicos, sobre relaciones y sobre clasificación de los datos. Un buen gráfico en sector muestra un compuesto total de partes (sectores) y cómo un sector afecta al total.

2. Diseño

- a. Dibujar primero un círculo de 360 grados; dado el hecho que 3,6 grados equivale a un 1%, un segmento de 72 grados representa un 20%. Cada segmento debe igualar su porcentaje en grados; para encontrar los grados, multiplicar el porcentaje del segmento por 3,6. El total de los segmentos del círculo debe sumar 360 grados, lo que equivale al 100%.
- b. Ubicar la primera línea radial a las «12 horas» en el gráfico circular, la cual es un buen punto de referencia para el lector.
- c. Partiendo de las 12 horas, arreglar los segmentos en el sentido del movimiento de los punteros del reloj en orden de tamaño decreciente, colocando el más grande primero. Si la lógica de la discusión impide el arreglo en orden de tamaño decreciente, asegurar que el sistema de ordenamiento —alfabético, cronológico u otro— queda claro para el lector.
- d. Utilizar, por lo menos, tres segmentos en el gráfico (el utilizar solamente dos resulta absurdo en un gráfico en sector) y no incluir más de cinco o seis. Si se necesita más de seis, entonces considerar un nuevo sistema de agrupación de la información con menos de 6 categorías o la utilización de otro tipo de gráfico

- e. No hacer segmentos más pequeños que 2% (7,2°). Si se tiene más de dos segmentos tan pequeños tratar de agruparlos juntos en uno más grande con el rótulo de «Otros», «Varios», o «Resto».
- f. Etiquetar todas las secciones horizontalmente. Usar paralelismo gramatical en las etiquetas a fin de crear continuidad. Incluir el porcentaje de cada segmento en su rótulo y tratar de rotular dentro del segmento. Si el segmento es demasiado pequeño para albergar un rótulo, ubicar el rótulo fuera del gráfico con una línea recta de flecha dibujada hacia dentro del gráfico.
- g. Desprender segmentos del gráfico para crear énfasis. Si se desprenden más de dos segmentos del círculo, éste pierde el efecto unificador.
- h. Ser cauteloso en el uso de colores en el gráfico circular. Los colores pueden aumentar la comprensión del gráfico circular, pero también pueden confundir al lector al distorsionar los datos. El crear demasiado énfasis puede confundir al lector hasta el punto de dudar sobre la veracidad del gráfico.

3. Ejemplos de gráficos en sectores

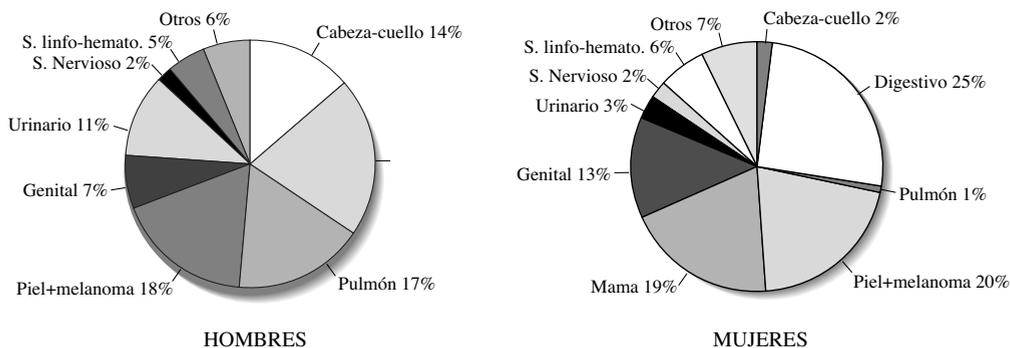


Gráfico IV.34. Incidencia de cáncer. Frecuencia relativa por órganos y aparatos. Granada 1988-90.

Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.34 es circular y muestra la incidencia de cáncer en Granada por órganos y aparatos, según sexo (nueve sectores).

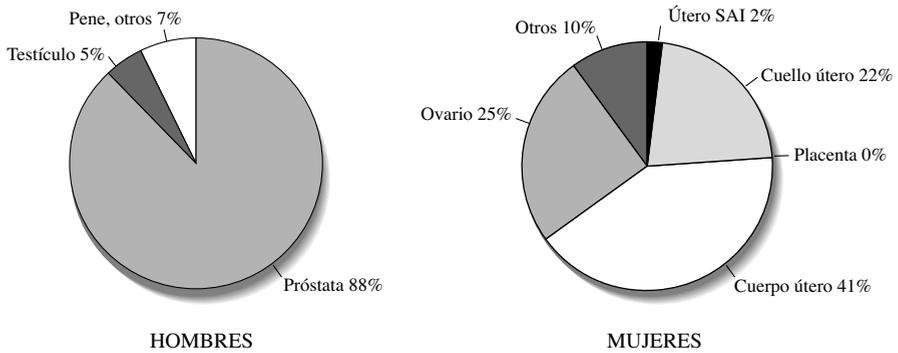


Gráfico IV.35. Incidencia de cáncer del aparato genital. Frecuencia relativa por localización y sexo. Granada 1988-90.
 Fuente: Martínez García, C. *El Cáncer en Granada, Incidencia y mortalidad 1988-90*. Granada: Escuela Andaluza de Salud Pública, 1994.

El Gráfico IV.35 es circular y muestra la incidencia de cáncer del aparato genital en Granada, por localización y sexo (tres sectores para hombres y cinco sectores para mujeres).

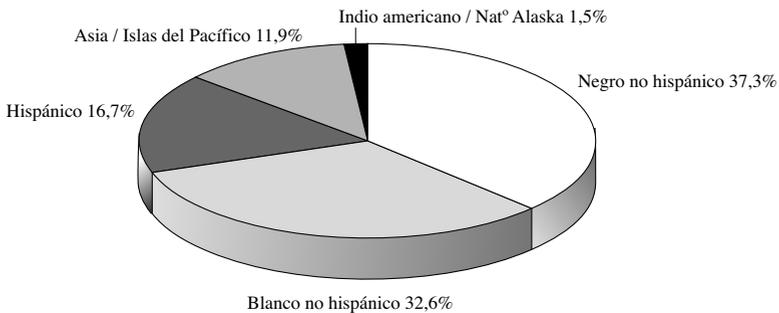


Gráfico IV.36. Porcentaje de casos de tuberculosis según raza y etnia. Estados Unidos, 1989 (N= 23.495).
 Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

El Gráfico IV.36 en sectores tridimensional el cual muestra el porcentaje de casos de tuberculosis según raza y etnia.

Gráficos de dispersión

También denominado *diagrama de dispersión*, es un gráfico que muestra la relación entre dos variables continuas, donde el eje X (horizontal) representa una variable y el eje Y (vertical) representa la otra variable.

Para crear un gráfico de dispersión debemos tener un par de valores para cada persona, grupo y otra entidad en nuestro conjunto de datos, un valor para cada una de las variables. Entonces diseñamos cada par de valores colocando un punto sobre el gráfico donde los dos valores se cruzan (ej. por ciento de sobrevivida y enfermedad arterial periférica; niveles de tetraclorodibenzeno-dioxin ajustado para lípidos y años de exposición; volumen expiratorio forzado y estatura; presión arterial sistólica y edad en mujeres; peso del niño al nacer y aumento de peso de la madre durante el embarazo).

Al interpretar un gráfico de dispersión miramos a la configuración total de los puntos marcados. Un modelo claramente compacto, indica un alto grado de correlación. Al contrario, un modelo con puntos muy esparcidos indican poca correlación.

1. Uso

El uso más frecuente de los gráficos de dispersión es en los estudios de correlación lineal de dos variables. En efecto, el paso inicial en cualquier problema de correlación -es decir la tendencia de una variable a relacionarse con los cambios de otra variable- es la construcción de un gráfico de dispersión. Este tipo de gráfico nos da una evidencia visual pronta de la presencia o ausencia de correlación; además nos indica si esa relación es positiva o negativa, y si es lineal o no lineal.

Ocasionalmente el investigador puede encontrar que el gráfico de dispersión es todo cuanto necesita para aprobar su punto. Esto es particularmente cierto cuando el investigador desea meramente señalar la presencia o la ausencia de correlación y cuando el gráfico de dispersión muestra una correlación muy alta o muy baja.

Cuando el investigador quiere avanzar un paso más y determinar el grado de asociación existente entre las dos variables, entonces puede medirlos mediante el uso del coeficiente de determinación, desde cero (no correlación) hasta 1.0 (correlación perfecta). Alternativamente puede usar el coeficiente de correlación -el cual es la raíz cuadrada del coeficiente de determinación- que varía desde 0 hasta + 1.0.

2. Diseño

- a. Se desarrolla una escala para las dos variables la cual se construye en el ángulo de la derecha, lo cual permite una gráfica rectangular de coordenadas.
- b. Cada variable se asigna a uno de los dos ejes de coordenadas; los valores de una variable (X) se sitúan en el eje horizontal; los valores de la segunda variable (Y) se sitúan en el eje vertical.
- c. Los datos se refieren a un número de sujetos, cada uno de los cuales provee observaciones sobre dos variables.

- d. Cada sujeto define un punto el cual corresponde a los valores observados de las dos variables.
- e. La dirección de la inclinación de los puntos indica la dirección de la correlación.
- f. La correlación es positiva cuando se asocian valores altos de la primera variable con valores altos de la segunda variable; la correlación es negativa (inversa) cuando se asocian valores altos de la primera variable con valores bajos de la segunda variable.
- g. El diagrama de dispersión permite ajustar o acomodar una línea recta a través de los puntos dispersos, mediante la técnica de la regresión lineal.

3. Ejemplos de gráficos de dispersión

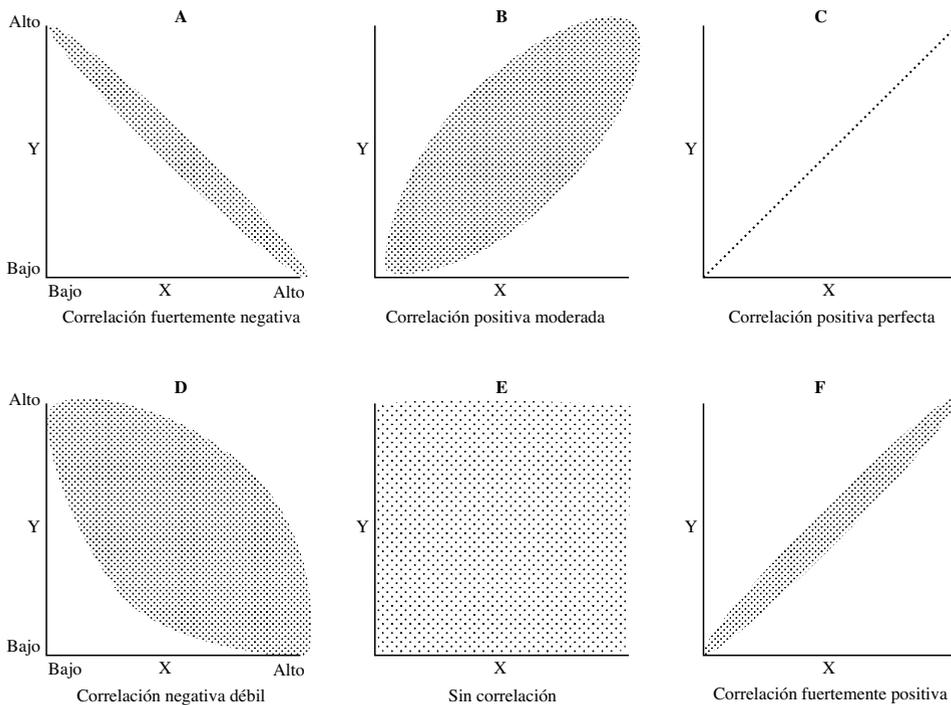


Gráfico IV.37. Seis tipos diversos de correlación.

Fuente: Elaboración propia.

El Gráfico IV.37 es un diagrama de dispersión el cual muestra seis tipos diversos de correlación: fuertemente negativa, negativa débil, sin correlación, positiva moderada, fuertemente positiva y positiva perfecta.

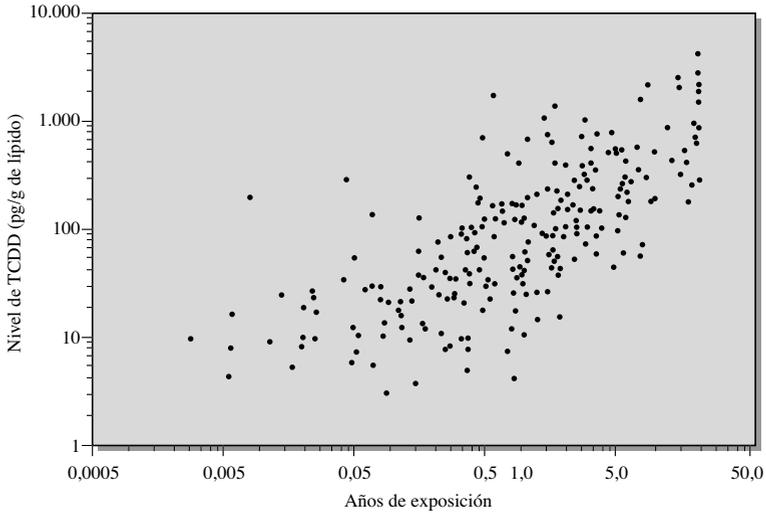


Gráfico IV.38. Niveles séricos de tetracloro dibenceno-P-Dioxin ajustado para lípidos según años de exposición. Estados Unidos, 1987.

Fuente: Centro para el control de enfermedades. *Informe de Toxi-infecciones alimenticias*. Atlanta: CCE, 1989.

El Gráfico IV.38 es un diagrama de dispersión que muestra los niveles de tetracloro-dibenceno-p-dioxin ajustado para lípidos en trabajadores según años de exposición.

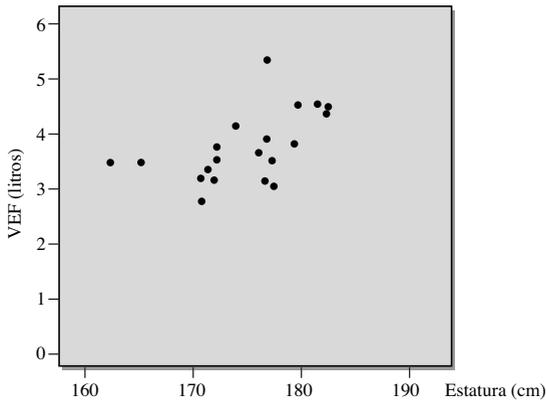


Gráfico IV.39. Relación entre volumen expiratorio forzado (VEF) y estatura en alumnos de medicina. Santiago, Chile, 1992.

Fuente: Universidad de Chile. Facultad de Medicina. *Volumen expiratorio forzado y estatura en alumnos de medicina*. Santiago: Editorial Universitaria, 1992.

El Gráfico IV.39 es un diagrama de dispersión el cual muestra la relación entre el Volumen Expiratorio Forzado (VEF) y la estatura en alumnos de medicina de sexo masculino. Nótese que la escala vertical Y (VEF) necesita el cero; la escala horizontal X (estatura) no precisa del valor cero, ya que la estatura de alumnos hombres de medicina va desde 160 a 190 centímetros. La estatura parece ser un factor predictor importante del VEF.

9. Notas a pie de página

Concepto y objetivos

Las notas a pie de página constituyen un excelente recurso descriptivo el cual se aprovecha escasamente por los investigadores clínicos. Las notas a pie de página introducen material aclaratorio o de documentación de naturaleza importante pero no esencial cuya inclusión en el texto quebraría su continuidad, desviando la atención del lector.

Las notas a pie de página son de dos clases: de referencia y de contenido. Las notas de referencia son aquellas que citan de las fuentes de los hechos que se publican o que hacen referencia a otras partes del documento mismo. Las notas de contenido, por otra parte, proveen información adicional, que el autor piensa que sería de interés para el lector pero que no es suficientemente vital para incorporarlo en el texto principal. Las notas de contenido son análogas a los apéndices. Aquellas tablas, esquemas, listas, cartas y otras similares las cuales no tienen una relevancia inmediata con el texto quedan mejor ubicadas en un apéndice; este material se refiere en el texto mediante una nota a pie de página. Cuando se trate de material simple y breve, se debe incluir directamente en la nota a pie de página, en vez de situarlo en un apéndice.²⁶

Usos

Las notas a pie de página tienen cuatro usos principales: a) Citar la fuente de afirmaciones en el texto que no califican como referencias propiamente tales; ésto abarca tesis, artículos en semanarios o periódicos, ponencias a jornadas o congresos no publicadas, cartas personales, opiniones específicas de personalidades, citas entre comillas, etc.; b) Hacer referencias cruzadas; c) Hacer comentarios incidentales o circunstanciales que amplían la discusión del texto lo cual podría beneficiar al lector pero que al introducirse en el texto interrumpirá su flujo de pensamiento; d) Hacer reconocimientos a personas o a instituciones.

El tipo de notas mencionadas en a) y b) corresponden a notas de referencia; las categorías c) y d) corresponden a notas de contenido. Una nota de contenido puede incluir uno o más asuntos.²⁷

Numeración

El lugar en el texto en el cual se introduce la nota se debe marcar con un número árabe ubicado un poco por encima de la línea. El número de la nota ha de seguir siempre al pasaje al cual se refiere. Si el pasaje es una cita entre comillas, el número de la nota viene al final de la cita. Los números de las notas deben seguir uno a otro en orden numérico, empezando con el número 1. La numeración debe empezar al comienzo de cada capítulo. Si el documento no está dividido en capítulos, la numeración correrá en forma consecutiva de principio a fin, teniendo cuidado de asegurar que la secuencia final es correcta. Debe evitarse el confundir la numeración de las notas a pie de página con la numeración correlativa de las referencias citadas en el texto. Para este efecto las últimas se indican con números árabes entre paréntesis.

Ubicación

Las notas deben disponerse en orden numérico al pie de la página o al final del capítulo. El uso de notas al pie de la página permite al lector leer el documento de corrido de principio a fin, sin la incomodidad de tener que interrumpir la lectura para buscar una referencia al final del capítulo o del documento y regresar nuevamente a la página y línea del texto original. Esto es particularmente engorroso cuando el documento se conserva en la forma de microficha.

La nota a pie de página debe empezar en la parte más baja de la página en que se la cita, aún cuando puede extenderse a la parte baja de la página siguiente si su extensión lo requiere. El texto de la nota debe separarse del texto principal a cuatro espacios con una línea continua de 3 cm. entre ambos.

Abreviaturas

El texto de la nota debe evitar las abreviaturas. Sin embargo las abreviaturas son de uso frecuente en las bibliografías, listas de referencias, materias tabulares y materiales ilustrativos.

No se deben utilizar abreviaturas que designan una parte de una palabra escrita. (Ej.: vol. por volumen; pt. por parte; cap. por capítulo, etc.) a menos que esa parte

de la palabra sea seguida por un número (Ej.: vol. 2; 4 vols.; pt.4; cap.10, etc.) Si no se acompañan de un número, esas palabras representadas por las abreviaturas deben escribirse completas.

Los títulos de revistas, diccionarios, enciclopedias, etc. pueden abreviarse utilizando las iniciales de las palabras de sus nombres; en el caso de las revistas biomédicas se aconseja seguir las abreviaturas utilizadas en el Index Medicus.

Los textos que utilizan notas a pie de página para referencias citadas en el texto en lugar de agrupar las referencias en una sección especial se ciñen a los mismos criterios y normas descritos en el Capítulo C, sección 13. REFERENCIAS de esta Parte IV de la obra.

Uso de «*IBIDEM*», «*OPERE CITATO*» y «*LOCO CITATO*»

Después de la primera mención completa de una referencia, menciones posteriores de la misma fuente se efectúan en forma abreviada. Esto se aplica solamente cuando las notas a pie de página se numeran consecutivamente. En consecuencia, las formas abreviadas no se pueden utilizar para referencias citadas en capítulos anteriores con numeración propia; si cada capítulo o parte de un texto numera las notas a pie de página en una serie propia, entonces una referencia citada en un capítulo previo debe citarse en forma completa en un capítulo siguiente.

La abreviación IBID, por la palabra latina *IBIDEM* «en el mismo lugar», se usa para repetir la referencia que precede cuando referencias al mismo trabajo se siguen sin ninguna otra referencia entremedio. Si la nueva mención es al mismo volumen y número de página, entonces se incluye solamente IBID. Si la nueva referencia al mismo trabajo es a un volumen o página distinta, se incluye IBID más los cambios.

La abreviación OP.CIT., por las palabras latinas *OPERE CITATO* «en el trabajo citado», se usa para una referencia previamente citada de forma completa pero que no la precede inmediatamente. En este caso se debe incluir solamente el apellido del autor seguido de OP. CIT más el número de la página, si éste fuera el caso.

La abreviación LOC. CIT., por las palabras latinas *LOCO CITATO* «en el lugar citado», se usa en lugar de IBID cuando la referencia al trabajo la precede inmediatamente y además se refiere a la misma página. LOC. CIT., también se usa en lugar de OP. CIT. cuando se hace referencia a un trabajo previamente citado y a la misma página. En consecuencia LOC. CIT. nunca es seguido por un volumen o un número de página.

10. Discusión

Concepto y objetivos

Así como el capítulo de resultados constituye el «corazón» del artículo, la discusión es la «coronación del proceso de la investigación» en su fase interpretativa. El capítulo de discusión responde a la pregunta: «¿Cuál es el significado de los hallazgos de la investigación?». El capítulo de discusión -en general- suele ser el más débil de todos los capítulos del artículo original. De hecho es el más vulnerable a críticas externas legítimas.

En el capítulo de resultados se muestran los hallazgos en forma clara y rigurosa. En la discusión el autor debe desagregar esa información mediante un análisis y reconstituirla como una entidad conceptual independiente mediante una síntesis y evaluación que le permita formular conclusiones.

El autor debe comentar los hallazgos más relevantes analizándolos de forma cautelosa. Naturalmente no tiene cabida aquí la repetición pesada e inútil de los hallazgos. La discusión debe centrarse primero en aquellos hechos de incuestionable importancia, para pasar luego a aquellos otros más controvertibles dónde cabe más de una interpretación legítima. En el segundo caso el autor debe ponderar las evidencias y los argumentos y proponer una explicación con fundamentos.²⁸

La forma como se organiza y se escribe este capítulo es una llave del éxito. El lector debe poder descubrir una secuencia lógica en el análisis de la información y sus implicaciones. Esta secuencia no sigue necesariamente el orden en que se condujeron los experimentos.

El capítulo de la discusión le otorga al autor una tribuna legítima para convencer al lector que la investigación tiene validez interna, es decir, que las diferencias observadas en los individuos o grupos participantes son atribuibles a la variable independiente (causa) y no a otras variables extrañas o al azar. Esto justifica los mecanismos de control del diseño (plan del estudio). Los experimentos verdaderos poseen alto grado de validez interna por el uso de procedimientos de control efectivos como la distribución aleatoria y el establecimiento de grupos de comparación. Los cuasi-experimentos carecen de asignación en forma aleatoria de los sujetos a las condiciones experimentales lo cual reduce su validez interna. Esta situación se hace aún más difícil o imposible en estudios con diseño no experimental (también llamado «observacional»).

Un aspecto que atenta gravemente contra la validez del estudio, tanto la validez interna como la externa, es el de los sesgos. Aquí destacan: a) Los sesgos de selec-

ción (criterios de inclusión y de exclusión en la muestra); b) Los sesgos de información (error sistemático en la medición de la variable); c) Los sesgos de confusión (error por presencia de otras variables con influencia sobre la variable dependiente).

Otro aspecto importante del capítulo de discusión es el de la validez externa, es decir el grado en que se pueden generalizar o extrapolar los resultados obtenidos de la muestra estudiada a la población de donde provino la muestra. La validez externa responde a la pregunta: ¿A qué población (o poblaciones), y en qué medios y condiciones pueden aplicarse los resultados del estudio? ²⁹

Conviene tener presente la distinción entre población accesible y población diana o población blanco. La primera es la población de sujetos que estuvo disponible para el estudio. La segunda es el grupo total de sujetos a quienes se podría generalizar -de manera lógica- los resultados. La generalización a la «población blanco» es muy riesgosa y no puede hacerse con la misma confianza que la generalización a la «población accesible».

La discusión debe relacionar los resultados con las teorías subyacentes y con los resultados de estudios afines. Por ello los hallazgos presentados en el capítulo de resultados se analizan y comparan con la información comunicada en publicaciones de otros autores. Esto significa, de hecho, la segunda revisión bibliográfica. (La primera revisión bibliográfica acaeció en la «Introducción» a fin de conocer los antecedentes del problema a investigar y configurar el marco teórico de la investigación).

La discusión debe analizar aquellos resultados inesperados y juzgar las posibles razones que los explican. También la discusión debe valorar el alcance de la investigación a la luz de las limitaciones del mismo, reubicando los hallazgos en el armazón del conocimiento científico al momento de la publicación. Esto exige un equilibrio entre la sobrevaloración y la subestimación de los hallazgos.

La discusión debe ser concisa, evitando redundancias, rodeos inútiles y consideraciones irrelevantes.

Este capítulo del artículo tiene por objeto:

- a. Examinar e interpretar los resultados en función de su evidencia y de sus limitaciones;
- b. Determinar la relación de coherencia o contradicción entre los datos presentados;
- c. Señalar las similitudes y las diferencias entre los resultados;

- d. Determinar la validez interna de la investigación;
- e. Determinar la validez externa de la investigación, haciendo generalizaciones (extrapolaciones) desde la muestra a la población accesible, a la población diana y a otras poblaciones;
- f. Destacar cualquier consecuencia teórica de los resultados;
- g. Sugerir mejoras en la investigación del tema;
- h. Señalar áreas distintas que hace falta investigar sobre el tema en cuestión.

Estructura

A fin de que el capítulo de discusión asegure los objetivos descritos anteriormente es necesario primero el comentar los propios resultados, tratando de desentrañar su significado. El análisis de los propios resultados facilita su interpretación en relación al estado del conocimiento específico.

Cumplida esta fase se deben redactar comentarios sobre los propios resultados comparándolos con los resultados publicados por otros autores. Esto obliga a una segunda revisión bibliográfica, con las citas correspondientes; no se trata de una revisión bibliográfica exhaustiva sino únicamente de las publicaciones más recientes y relevantes que tengan relación con el problema estudiado y los hallazgos presentados.

La secuencia debe ayudar a la comprensión del análisis interpretativo de los resultados. Una posibilidad es seguir una secuencia más o menos paralela a la descripción de los resultados, pero con la libertad y la flexibilidad que el caso aconseje.

La discusión, etapa especulativa del comunicado de la investigación, debe poder concretar una respuesta (conclusión) a la pregunta de investigación planteada en la introducción en la forma de objetivo o hipótesis. Surgen aquí tres posibilidades: a) Hay respuesta, la investigación logró resolver el problema de investigación; b) No hay respuesta del todo; c) Hay respuesta parcial; el estudio resolvió algunos elementos del problema de investigación; Aquí caben sugerencias concretas para nuevas investigaciones.

Estilo

El estilo descriptivo y narrativo de los capítulos anteriores cambia a un estilo argumentativo, cuya finalidad es convencer al lector de que los resultados son válidos,

son importantes, son novedosos y son útiles. El estilo argumentativo incluye elementos de debate y de controversia.

La redacción de la discusión constituye un desafío a la capacidad del autor para diferenciar lo que parece cierto frente a lo que parece probable y a lo que parece meramente tentativo; un desafío a su capacidad de objetividad, de ponderación, de cautela, de prudencia y de honestidad intelectual.

El tono del lenguaje debe ser seguro, sin llegar a ser «triumfalista». Las afirmaciones no pueden ser siempre categóricas frente al relativismo del conocimiento adquirido, el cual suele trabajar más con relaciones que con valores aislados. Problemas no resueltos tales como la representatividad de la muestra, el control de las variables, la validez de los instrumentos de medición, la confiabilidad de las fuentes de información -tanto documentales como de campo- y la probabilidad de error en el proceso de inferencia, cobran un alto precio en este capítulo.

Faltas frecuentes

- a. Se vuelve a formular conceptos e informaciones ya planteados en la introducción.
- b. Se repite la información de los resultados.
- c. No se confrontan los resultados.
- d. Se reformulan los mismos puntos ya tratados.
- e. Se polemiza en forma trivial.
- f. Se hacen comparaciones teóricas sin fundamento.
- g. Se hacen conjeturas sin identificarlas como tales.
- h. Se hacen planteamientos sin relacionarlos lógicamente con la información empírica o teórica disponible.
- i. No se comentan todos los resultados importantes.
- j. No se interpreta cada resultado en términos de la hipótesis original con la que se relaciona.
- k. No se interpreta cada resultado a la luz de hallazgos de estudios semejantes de otros autores.
- l. Las interpretaciones no dan debida consideración a las limitaciones del estudio.
- m. Las interpretaciones no son congruentes con los resultados.
- n. Se hacen interpretaciones de causalidad no justificadas, cuando de hecho se trata de simples asociaciones.

- o. Las interpretaciones no se organizan en una forma significativa.
- p. Hay sospechas o evidencias de errores y/o de sesgos sistemáticos en las interpretaciones.
- q. Se hacen generalizaciones que no se justifican con base en la muestra utilizada.
- r. No se hacen recomendaciones respecto a la forma de mejorar los métodos de estudio.
- s. No se hacen recomendaciones para futuras investigaciones.

11. Conclusión

Concepto y objetivos

La última parte del cuerpo del artículo científico original se denomina conclusión. De hecho la conclusión es la deducción, consecuencia o resolución que el autor propone al lector, luego de un completo razonar en la discusión. La conclusión es el fin o terminación del artículo científico original. Concluir, del latín «conclusio» es inferir o deducir una verdad de otras que se admiten, presuponen o demuestran.

La conclusión del artículo halla su homólogo en la peroración de un discurso, en la que se hace la enumeración de las pruebas y se trata de mover con más eficacia que antes el ánimo del auditorio mediante el reforzamiento de las impresiones causadas presentando la causa desde el punto de vista más favorable, recapitulando las principales razones y moviendo los afectos.³⁰ La conclusión del artículo tiene su equivalente en el epílogo, última parte de una composición literaria, sea dramática o novelística, en la que se recapitula lo que antecede.

La conclusión representa la *respuesta* de la investigación a la *pregunta* planteada al inicio de la misma, en la forma de objetivo o hipótesis. En la conclusión se expresa en forma sucinta la respuesta precisa a la pregunta planteada en la «Introducción» sea que ésta coincida o no con la hipótesis propuesta. La conclusión también responde a las interrogantes que condujeron al diseño y a la realización de la investigación.

En la conclusión se pueden añadir las deducciones que surgen de la confrontación de los propios datos con los datos publicados por otros autores, hecha en la «Discusión». La conclusión debe apoyarse enteramente en los resultados de la investigación en cuestión, mediante un juicio objetivo e imparcial. En consecuencia no tienen cabida en la conclusión juicios de valor del autor sin base en los hallazgos, aún cuando parezcan lógicos y de sentido común.³¹

Estructura

La conclusión se ubica habitualmente como un párrafo aparte al final del capítulo de «Discusión». Rara vez aparece como un capítulo aparte en los artículos originales. La extensión de la conclusión es habitualmente breve y debe relacionarse directamente con los objetivos o hipótesis. Si hubo más de un objetivo, habrá más de una conclusión, sea esta positiva o negativa.

Estilo

El estilo debe ser claro y directo, en tono afirmativo; de hecho se debe evitar el estilo ampuloso, farragoso, por cuanto le resta claridad al mensaje. Además debe ser conciso evitándose todo exceso de ideas y palabras innecesarias. El tono no debe ser «triumfalista»; no debe ir más allá de lo que los hechos permiten. El estilo debe tener un equilibrio entre imaginación y objetividad, entre entusiasmo y reflexión.

Faltas frecuentes

- a. El artículo no tiene conclusiones explícitas, lo cual deja sin respuesta la pregunta de la investigación.
- b. Las conclusiones no se justifican por cuanto no se apoyan en los hallazgos del estudio, aun cuando la conclusión aparezca razonable y enmarcada dentro del sentido común.
- c. Las conclusiones no concuerdan con los objetivos e hipótesis (preguntas de investigación).
- d. Las conclusiones no se entienden. El estilo es poco claro, oscuro, ampuloso, farragoso. El estilo no es conciso; hay conceptos sobrantes y palabras sobrantes.

12. Agradecimientos

Concepto y objetivos

La realización de una investigación requiere, con frecuencia, de la cooperación de instituciones y de personas que no califican como autores. Estas ayudas suelen suceder en la generación de ideas, en la selección del problema, en la elección del diseño adecuado, en el tratamiento estadístico, etc. Otras veces los autores reciben una cooperación técnica concreta tal como la provisión o facilitación de equipos o materiales necesarios. También suele acontecer alguna ayuda financiera directa al proyecto de investigación en forma de alguna subvención, de becas, contratos, etc. También conviene mencionar la cooperación de quienes leyeron críticamente el manuscrito y aportaron mejoras significativas a su contenido, su estructura y su estilo.⁶

En los agradecimientos el autor reconoce a los mentores y colegas, enumera a los individuos o instituciones que apoyaron la investigación y le da crédito a los trabajos citados en el texto. No parece apropiado aquí el dar reconocimiento a personas que colaboraron dentro de la rutina diaria de su trabajo. El reconocimiento debe reservarse para aquellos cooperadores que lo merezcan de verdad. En consecuencia esta parte del artículo tiene por objeto: a) Reconocer la cooperación de personas o instituciones que ayudaron en forma intelectual o material al autor en su investigación; b) Reconocer la cooperación de personas que ayudaron en la redacción del artículo; c) Reconocer la cooperación especial de personas que revisaron el manuscrito tanto en aspectos de su contenido como de su estructura y de su estilo.

El autor debe resistir la tentación -y a veces hasta la presión- de incluir a algunos de estos colaboradores en la lista de autores. A este respecto valen los criterios de autoría descritos en el capítulo C.2. AUTORES.

Estructura

Los agradecimientos suelen ubicarse en una sección de «Agradecimientos» inmediatamente después de la «Conclusión» y antes de la sección de «Referencias». En otros tipos de publicaciones los agradecimientos se ubican al final de la introducción o en una nota a pie de página.

Estilo

La redacción de este capítulo tiene poco o nada que ver con lo científico. El eje de la redacción aquí se encuentra en la atención, el respeto y el afecto que tiene una persona a otra: en la verdadera cortesía. En consecuencia los elementos que caracterizan esta sección son la verdad, la naturalidad, la sobriedad y la brevedad. No se debe pecar por defecto ni por exceso; tampoco cabe aquí la vaguedad y la redacción debe ser directa y específica.

Faltas frecuentes

Las faltas más frecuentes caen en las siguientes tres categorías:

- a. Agradecer la contribución de personas que bien podrían calificar como coautores del trabajo.
- b. No reconocer la cooperación de personas o instituciones que claramente hicieron posible la investigación.
- c. Faltar a la ponderación en la exposición del agradecimiento.

13. Referencias

Concepto y objetivos

Etimológicamente referencia proviene del latín «referens», «referentis», referente. En el artículo original la referencia es la indicación en el escrito de la forma y el lugar del artículo al que se remite al lector.³²

El formato de las citas bibliográficas es un asunto que ha preocupado a la comunidad científica y a las revistas biomédicas en los últimos 20 años. Esto ha sido el resultado de las quejas de los autores sobre la multitud de estilos en la presentación de las citas que había en uso en las diversas revistas, lo que obligaba a reescribirlas cada vez que una revista rechazaba un artículo y el autor quería enviarlo a otra.

Un criterio fundamental de inclusión de referencias es que el lector pueda recuperarlas para consulta y verificación.

El «Comité Internacional de Editores de Revistas Médicas» ha ido imponiendo, por consenso de los editores de las revistas médicas más importantes del mundo, el uso de la modalidad de «Orden de Mención». Este consiste en que a cada referencia se le otorga un número consecutivo según su orden de aparición en el texto. El formato de las referencias que utiliza este sistema -también llamado «Sistema Vancouver»- se basa en el que utiliza la Biblioteca Nacional de Medicina (National Library of Medicine) de los Estados Unidos en el *Index Medicus*.¹¹

El capítulo de referencias tiene los siguientes objetivos:

- a. Dar a conocer trabajos anteriormente publicados para apoyar o refutar la información presentada. Esto implica el identificar las fuentes originales de ideas, conceptos, métodos, técnicas y resultados provenientes de estudios anteriores publicados;
- b. Prestar fiabilidad a la información documentando su origen. Dar solidez a los hechos y opiniones expresadas por el autor;
- c. Permitir que el lector se informe más a fondo sobre algún punto de su interés. Orientar al lector para que se informe en mayor extensión y profundidad sobre aspectos relevantes del estudio;
- d. Cumplir con el principio ético de no plagiar material ajeno.

Las referencias citadas deben cumplir con los siguientes requisitos:

- a. Citar solo trabajos publicados que se consideran importantes para el desarrollo de las ideas de la investigación en cuestión;
- b. Citar trabajos que el lector pueda obtener;
- c. Dar, para cada referencia, una descripción suficiente del trabajo para que cualquier lector la pueda encontrar y consultar;
- d. Citar nombres y títulos tal como aparecen en la publicación original, incluso con la misma puntuación. Cabe señalar que el título completo de los libros aparece muchas veces en la portada interior, no en la cubierta;
- e. No son referencias aceptables datos, tesis, cartas y discursos no publicados, obras en prensa excepto aquellas formalmente aceptadas y con fecha tentativa de publicación, comunicaciones verbales y resúmenes. Éstas deben citarse a pie de página. Aun así, deben ser citadas en forma completa;
- f. Citar en el texto, de acuerdo con el número asignado, cada trabajo del capítulo de referencias;
- g. Asignar un solo número a cada referencia y repetir el mismo número cada vez que se menciona la misma referencia;
- h. Ubicar la sección de referencias en folio separado al final del trabajo;
- i. Seguir las normas de la revista a la cual se somete el artículo para publicación;
- j. De no estar seguro de esas normas, se debe dar toda la información bibliográfica para que los editores escojan los datos que interesan, según el formato de la revista particular.

Aunque en una revista se usen las abreviaturas de publicaciones periódicas indicadas, los autores deben dar los títulos completos de las revistas. Cuando se trata de una publicación poco conocida a escala internacional, es preciso indicar la ciudad y país donde se publica.

Es absolutamente imprescindible cotejar varias veces el texto terminado contra las referencias para asegurar que se han citado correctamente, especialmente cuando otra persona ha mecanografiado el manuscrito.

Es aconsejable el conservar todos los datos de las referencias hasta que se haya publicado el artículo, cada una en una tarjeta aparte o en la memoria de un ordenador.

En un artículo científico original las referencias suelen encontrarse principalmente en el capítulo de «Introducción» (primera revisión bibliográfica) y en el capítulo de «Discusión» (segunda revisión bibliográfica). Ocasionalmente se encuentran referencias en el capítulo de «Materiales y Métodos».

Estructura

Los sistemas de citación y referencia más comunes son:

- a. «*Orden de mención*»: Citar por números consecutivos (entre paréntesis, en «superíndice» o «volado») y enumerar las referencias en el mismo orden de mención.
- b. «*Nombre y año*» (estilo de Harvard): Hacer una lista alfabética de referencias (sin números) y citarlas por el apellido del autor y año de publicación.
- c. «*Numérico-alfabético*»: Enumerar las referencias en orden alfabético por los apellidos de los autores. En ese caso, los números en el texto no son consecutivos.³³

a. Ejemplo de texto y referencias por «Orden de mención»:

Texto

«Es un hecho bien conocido que las tasas de incidencia y de mortalidad por cáncer de pulmón en los países desarrollados son más altas que en la mayoría de los países de América Latina y el Caribe (1,2). En cambio, en América Latina son más elevadas las de cánceres de vejiga urinaria, laringe y cavidad bucal (1,3). En dos estudios realizados en Antioquia, estado del noroeste de Colombia, se informa de una alta frecuencia de cáncer de laringe, especialmente entre las mujeres (4,5)».

Referencias:

1. Puffer RR, Griffith GW. *Características de la mortalidad urbana. Informe de la Investigación Interamericana de Mortalidad*. Washington DC: Organización Panamericana de la Salud, 1968. Publicación Científica 151.
2. Zavala D, Correa P. Latin America. In: Howe GM, ed. *Global GeoCancerology. A World Geography of Human Cancer*. Londres: Churchill Livingstone, 1986, p. 98.
3. Waterhouse J, Muir C, Shanmugaratnam K, Powell J. eds. *Cancer Incidence in Five Continents*, vol. IV. Lyon, Centro Internacional de Investigaciones sobre el Cáncer, 1982. Publicación Científica 42.
4. Correa P. Statistical study of cancer in Antioquia. *Schweiz Z Pathol* 1955, 18:491-500.
5. De Restrepo HE, Franco A. Cancer morbidity and mortality among the insured population of the Social Security Institute of Antioquia, Colombia. *Int J Epidemiol* 1978, 7(3):285-291.

Ventajas:

- Permite una lectura rápida que puede interrumpirse para consultar las referencias citadas.
- Ahorra espacio a la revista.
- Concuerda con el sistema de Vancouver adoptado por varios centenares de revistas biomédicas importantes de alcance internacional.

Desventajas:

- Es laborioso añadir nuevas referencias, ya que hay que reenumerar las demás.
- Puede separar los distintos trabajos de un mismo autor.

b. Ejemplo de texto y referencias por «Nombre y año»:

Texto

«Es un hecho bien conocido que las tasas de incidencia y de mortalidad por cáncer de pulmón en los países desarrollados son más altas que en la mayoría de los países de América Latina y el Caribe (Puffer y Griffith, 1968; Zavala y Correa, 1986). En cambio, en América Latina son más elevadas las de cánceres de vejiga urinaria, laringe y cavidad bucal (Puffer y Griffith, 1968; Waterhouse et al., 1982). En dos estudios realizados en Antioquia, estado del noroeste de Colombia, se informa de una alta frecuencia de cáncer de laringe, especialmente entre las mujeres (Correa, 1955; De Restrepo y Franco, 1978)».

Referencias:

1. Correa P. Statistical study of cancer in Antioquia. *Schweiz Z Pathol* 1955, 18:491-500.
2. De Restrepo HE, Franco A. Cancer morbidity and mortality among the insured population of the Social Security Institute of Antioquia, Colombia. *Int J Epidemiol* 1978, 7(3):285-291.
3. Puffer RR, Griffith GW. *Características de la mortalidad urbana. Informe de la Investigación Interamericana de Mortalidad*. Washington DC: Organización Panamericana de la Salud, 1968. Publicación Científica 151.
4. Waterhouse J, Muir C, Shanmugaratnam K, Powell J. eds. *Cancer Incidence in Five Continents*, vol. IV. Lyon, Centro Internacional de Investigaciones sobre el Cáncer, 1982. Publicación Científica 42.
4. Zavala D, Correa P. Latin America. In: Howe GM, ed. *Global GeoCancerology. A World Geography of Human Cancer*. Londres: Churchill Livingstone, 1986, p.98.

Ventajas:

- Pueden añadirse o suprimirse fácilmente las referencias.
- Permite relacionar inmediatamente a personas con los fenómenos mencionados.

Desventajas:

- Es inconveniente para la lectura, especialmente cuando se citan varias referencias en un párrafo.
- Es más costoso para la revista por cuanto desperdicia espacio.

c. Ejemplo de texto y referencias «Numérico-alfabéticas».

Texto

«Es un hecho bien conocido que las tasas de incidencia y de mortalidad por cáncer de pulmón en los países desarrollados son más altas que en la mayoría de los países de América Latina y el Caribe (3,5). En cambio, en América Latina son más elevadas las de cánceres de vejiga urinaria, laringe y cavidad bucal (3,4). En dos estudios realizados en Antioquia, estado del noroeste de Colombia, se informa de una alta frecuencia de cáncer de laringe, especialmente entre las mujeres (1,2)».

Referencias:

1. Correa P. Statistical study of cancer in Antioquia. *Schweiz Z Pathol* 1955, 18:491-500.
2. De Restrepo HE, Franco A. Cancer morbidity and mortality among the insured population of the Social Security Institute of Antioquia, Colombia. *Int J Epidemiol* 1978, 7(3):285-291.
3. Puffer RR, Griffith GW. *Características de la mortalidad urbana. Informe de la Investigación Interamericana de Mortalidad*. Washington DC: Organización Panamericana de la Salud, 1968. Publicación Científica 151.
4. Waterhouse J, Muir C, Shanmugaratnam K, Powell J. eds. *Cancer Incidence in Five Continents*, vol. IV. Lyon, Centro Internacional de Investigaciones sobre el Cáncer, 1982. Publicación Científica 42.
5. Zavala D, Correa P. Latin America. In: Howe GM, ed. *Global GeoCancerology. A World Geography of Human Cancer*. Londres: Churchill Livingstone, 1986, p. 98.

Ventajas:

- Permite rápida lectura y ahorra espacio.
- Permite apreciar el conjunto de trabajos de un mismo autor.
- Es fácil de preparar para los autores.
- Está respaldado por el Consejo de Editores Biológicos (Council of Biology Editors) de los Estados Unidos.

Desventajas:

- Siempre hay que reenumerar las llamadas en el texto si se añaden nuevas referencias.
- Su ámbito de uso se restringe a revistas de biología; no se utiliza en las revistas médicas y de salud.

Estilo

La utilización de publicaciones, tanto para apoyar como para refutar la información presentada por diversos autores, se debe acreditar.³⁴

Las referencias de las obras publicadas han de ser lo más breves posibles en el texto; los datos bibliográficos completos figuran aparte, ya sea en un listado al final del trabajo bajo el título de «Referencias» o como «Notas a pie de página».

Existen varios tipos de referencias bibliográficas tal como se describió en la sección de «Estructura». Conviene elegir uno y atenerse a él, nunca mezclarlos en un mismo artículo. En el sistema «Orden de mención» cada referencia aparece en el orden de su primera mención en el texto y es numerada en orden consecutivo. En el sistema de «Nombre-año», las entradas se ordenan alfabéticamente por el nombre del autor, y cuando un autor (o autores) tiene dos o más entradas, éstas se ordenan cronológicamente.

Adviértase que en ambos estilos se subrayan los títulos de los libros y los títulos de las revistas. No se subraya el título del artículo de la revista ni el título del capítulo del libro.

Aunque cada revista tiene sus propias normas la mayoría se atiene a las normas del Comité Internacional de Editores de Revistas Médicas, versión 1993. A continuación se detalla la información mínima necesaria para poder localizar un documento y que es requerida por las revistas o casas editoriales. Si la revista opta, por ej., por no usar cursivas, no tiene más que modificar nuestra bibliografía. En cambio si le faltan datos, debería devolvernos el manuscrito para que lo complementemos.

EL ESTILO DE ORDEN DE MENCIÓN (O NUMÉRICO)

El sistema de citación y referencia de uso más frecuente en medicina y ciencias de la salud es el de «orden de mención», también denominado «estilo numérico». Es el sistema adoptado por las normas del Comité Internacional de Editores de Revistas Médicas, en 1988 y se basa en el formato utilizado en el *Index Medicus*.¹¹

En esta modalidad las referencias o citas de obras se indican en el lugar apropiado del texto -mediante números consecutivos- bien ligeramente por encima del texto (opción Nota de pie en WordPerfect), o bien al mismo nivel entre paréntesis. Es preferible que las citas estén al final de una frase. Ejemplos:

El profesor Watson hace un análisis comparativo de los distintos métodos contraceptivos en su obra *Ginecología*. 1981(1).

El profesor Winston recomienda prohibir el tabaco a cualquier edad (2).

Las citas literales se escriben entre comillas.

Las referencias bibliográficas completas irán numeradas como en el texto. Si se ofrecen juntas en una bibliografía al final, irán ordenadas por número:

1. Organización Mundial de la Salud. *Les infections respiratoires de l'enfant: leur traitement dans les petits hopitaux*. Ginebra: OMS, 1988.
2. Cerezo Jiménez MA, López Pérez P. *Prevalencia y duración de la lactancia materna en Andalucía*. Sevilla: Junta de Andalucía, 1986.
3. Junta de Andalucía. *Datos estadísticos básicos de la población andaluza: padrón, 1986*. Sevilla: Junta de Andalucía, 1989.
4. Casado de Frías E. *Lactancia natural*. Madrid: Ministerio de Sanidad y Consumo, 1983.

La lista de referencias al final de un artículo o informe debe indicar las páginas si se ha hecho referencia a las mismas en el transcurso del trabajo, o si se han transcrito citas literales.

Libros y monografías:

Apellido del autor(es) (o editores, compiladores, etc.), inicial (es) del nombre*. Título con subtítulo si lo hay. Mención de edición. Lugar: entidad, editorial y año.

Belle D, ed. *Lives in stress: women and depression*. Beverley Hills: Sage, 1982.

Capítulo de un libro:

Apellido del autor(es) del capítulo, inicial(es) del nombre del autor(es) del capítulo. Título del capítulo. En: Apellido del autor(es) del libro, iniciales del nombre del autor(es) del libro. *Título del libro*. Mención de edición. Lugar de publicación: entidad, editorial y año. Cuando proceda, se agregarán los números del volumen y páginas consultadas, serie a que pertenece y número que le corresponde en la serie.

de Oleaga Usategui JI, de Manuel Keenoy E**. Sistema de Salud y Atención Primaria. En: Martín Zurro A, Cano Pérez JF. *Manual de Atención Primaria. Organización y pautas de actuación en la consulta*. Segunda edición. Barcelona: Ediciones Doyma, 1989:12-28.

Artículo de revista:

Apellido del autor(es)*. Inicial(es) del nombre del autor. Título del artículo. *Título de la revista*. Año; vol: (nº), páginas.

Mckween RE. Resuscitation of the newborn. *Clin. Obstet. Gynecol.* 1987; 30(3), 611-20.

* Las normas del Comité Internacional de Editores de Revistas Médicas recomiendan mencionar a todos los autores. Cuando haya seis o más se debe citar solamente a los tres primeros y añadir *et. al.*

** En unos apellidos compuestos se puede preferir la transposición de los prefijos o partículas: el español "de", el francés "de" o "d", el holandés "van" y el alemán "von". Si se efectúa la inversión de los nombres compuestos en el ejemplo mencionado, esto aparecería así: Oleaga Usategua JI de, Manuel Keenoy E de.

EL ESTILO DE NOMBRE-AÑO (O DE HARVARD)

Es el sistema adoptado en historia, filosofía y ciencias sociales. En esta modalidad se inserta en el texto el apellido del autor citado y el año de publicación de su obra. Uno o ambos irán entre paréntesis dependiendo de la construcción de la frase:

Rodríguez (1989) comprobó este mismo fenómeno cuando realizó una encuesta de salud en un barrio valenciano; un estudio posterior (Mann, 1982) reveló que...

Si se desea hacer referencia a una página específica, es preferible hacerlo en el texto mismo en vez de hacerlo en la lista de referencias:

Colditz (1988, pág. 937) ha demostrado que las mujeres fumadoras tienen un mayor riesgo...

Las referencias completas se ofrecen en un listado al final del trabajo, ordenado alfabéticamente.

En el caso de que no haya autor, se sustituye por el título o por el nombre de la organización responsable de la obra:

Iavicoli, R. 1977. *Scuola e Salute*. Florencia: La nuova Italia.
Organización Mundial de la Salud. 1985. *Having a baby in Europe*. (Public Health in Europe, n° 26). Copenhague: OMS.
Valoración del estado de salud materno-infantil. 1987. Sevilla: Junta de Andalucía.

Los datos que se consignan en una referencia son los siguientes:

Libro:

Apellido del autor(es), iniciales del nombre del autor (o autores). Año de publicación. *Título: subtítulo*. Número de edición (si consta). Nombre y número de la serie (si es del caso). Lugar de publicación: editorial.

Pilliteri, A. 1976. *Maternal-newborn nursing: care of the growing family*. 2nd ed. Boston: Little, Brown.

Si está firmado por dos o tres autores, sus apellidos e iniciales se separarán mediante punto y coma:

Greene, W.H.; Simmons, P.; Morton, B.G. 1988. *Educación para la salud*. México: Interamericana.

Si los autores son más de tres, puede consignarse solamente el apellido y las iniciales del nombre del primero, seguidos de la abreviatura internacional *et al* (y otros): Año de publicación. *Título: subtítulo*. Número de edición (si consta). Nombre y número de la serie (si es del caso). Lugar de publicación: editorial.

Casado de Frías, E. *et al.* 1983. *Lactancia natural*. (Serie Monográfica; n° 7). Madrid: Ministerio de Sanidad y Consumo.

Si la lista incluye más de una obra del mismo autor publicada en el mismo año, éstas se distinguen añadiendo una letra tras el año. En la bibliografía al final no se repite el nombre del autor, sino que se pone un guión largo.

Si el autor es una institución:

Organización Mundial de la Salud, 1987. *Trastornos hipertensivos del embarazo*. (Serie de Informes Técnicos; n° 758). Ginebra: OMS.

En el caso de obras sin firma, la primera palabra del título, que no sea un artículo, ocupa el lugar del autor:

Collins pocket dictionary of the English Language. 1981. London: Collins.

Capítulo de libro:

Apellido del autor(es) del capítulo; Iniciales de los nombres del autor del capítulo. Año. Título del capítulo. *En* (seguido de dos puntos) Autor o encabezamiento alternativo del libro. *Título del libro*. Lugar; editorial, páginas del capítulo (inicial y final inclusive).

Want, P.C. 1987. *Obstetrics and Gynaecology En: Information sources in the Medical Sciences*. 3rd. ed. Eds. L.T. Morton and S. Godbolt. London, Butterworths.

Artículo de revista:

Apellido del autor(es). Iniciales de los nombre del autor. Título del artículo. *Título de la revista*, año; vol. (n°), págs. (inicial y final inclusive).

Kramer MS. Intrauterinic growth and retardaton. *Pediatrics*, 1987. 80(4): 502-11.

Faltas frecuentes

- a. No se presentan referencias.
- b. Las referencias no parecen pertinentes al tema del estudio.
- c. Las referencias se citan equivocadamente.
- d. Las referencias son demasiado exiguas.
- e. Las referencias son muy numerosas (sospecha de pseudoerudición del autor).
- f. Las referencias son obsoletas (sospecha de desactualización del autor).
- g. Las referencias no se citan en el texto.

- h. Se incluyen referencias al parecer no consultadas personalmente por el autor.
- i. Los documentos correspondientes a las referencias no son accesibles al lector.
- j. Las referencias citadas están incompletas:
 - (1) En el caso de libros y monografías falta información sobre apellidos e iniciales de los nombres de los autores, título, número de la edición, lugar, entidad, editorial y año.
 - (2) En el caso de capítulos de libros, no queda claro el título del capítulo frente al título del libro; tampoco aparece claro el nombre del autor del capítulo frente al autor -o editor principal- del libro.
 - (3) En el caso de artículos de revistas, falta información sobre apellidos e iniciales de los nombres de todos los autores, título exacto del artículo, nombre abreviado o completo de la revista, año, volumen, número del ejemplar, páginas inicial y final y año.
- k. Se incluyen como referencias documentos no aceptables para la comunidad científica. Tal es el caso de:
 - (1) Trabajos presentados a conferencias pero no publicados;
 - (2) Tesis de graduación inéditas;
 - (3) Manuscritos de artículos originales o de revisión presentados para publicación a revistas que se hallan en proceso de revisión y decisión sobre su publicación.
 - (4) Comunicaciones personales.
 - (5) Resúmenes.
 - (6) Artículos o editoriales de periódicos.
- l. La lista de referencias no se enumera consecutivamente según el orden de aparición de las citas en el texto, en el caso de utilizarse el estilo de «orden de mención» o «numérico».

14. Apéndices

Concepto y objetivos

El término «apéndice» se usa en escritura científica siguiendo el significado de su etimología latina «appendix»: cosa adjunta o añadida a otra, de la cual es como parte accesoria o dependiente.³²

Un apéndice, aunque no es de manera alguna una parte esencial del artículo original, es un instrumento útil para poner a la disposición del lector un material relacionado con el texto pero inapropiado para su inclusión en él. Los apéndices pue-

den contener, por ejemplo, tablas demasiado detalladas para su inclusión en el texto; grupos de ilustraciones; notas técnicas sobre los métodos; listas; cuestionarios utilizadas en la recolección de la información; estudios de casos demasiado extensos para incluirlos en el texto.

Los artículos científicos originales publicados rara vez incluyen un apéndice. La razón principal es la escasez de espacio en las revistas de prestigio, las cuales tratan de acomodar el mayor número posible de contribuciones originales. Si un lector está interesado en mayores detalles, entonces los solicitará por carta personal al autor principal.

El capítulo de «Apéndice» del artículo tiene por objeto el complementar o ilustrar el desarrollo del tema. Se aceptan dos criterios de inclusión de apéndices: a) Incluir información relevante que por su extensión o configuración no encuadra bien dentro del cuerpo del artículo; (Ej.: la copia de un cuestionario); b) Incluir información conveniente, a pesar de ser de importancia secundaria, para comprobación de datos. (Ej.: la fotocopia de datos en un anuario estadístico)

Estructura

El «Apéndice» debe ubicarse al final del artículo en folio separado, después de las referencias. Nunca debe ubicarse al final de los capítulos donde se lo menciona. Materiales de diferentes categorías deben ubicarse en apéndices separados. Si hay más de un apéndice, cada uno debe tener un número o una letra distinta.

Cada apéndice debe llevar un título descriptivo, el cual debe aparecer en el índice de contenido. Con respecto al espaciamiento, no es mandatoria la uniformidad; el texto de un apéndice puede estar a espacio sencillo y el texto de otro apéndice puede ir a doble espacio, dependiendo de la naturaleza y origen del material.

Estilo

Cada «Apéndice» conserva el estilo propio de su origen.

Se debe citar la fuente de la información cuando el apéndice no ha sido elaborado por el autor del artículo científico original.

Cuando aparecen documentos fotocopiados, tales como artículos publicados previamente, facsímiles de manuscritos, cuestionarios, etc. se debe agregar un número de página a cada hoja fotocopiada utilizando numerales árabes entre corchetes en la esquina superior derecha de cada página. Los corchetes indican que el número de la página no forma parte del documento principal.

Faltas frecuentes

Las faltas más frecuentes son:

- a. Incluir como «Apéndice» datos tabulados complementarios que el autor recopiló durante la investigación y que no son esenciales para la comprensión del mensaje;
- b. Incluir información complementaria -pero no esencial- en el cuerpo del artículo, haciendo caso omiso del recurso «Apéndice»;
- c. Recopilar la información en el apéndice de forma desordenada, dificultando así su lectura y comprensión.

D. PARTES DEL ARTÍCULO DE REVISIÓN

M. A. Bobenrieth Astete

1. Concepto y estándares

Las revisiones de la literatura publicadas en la forma de «artículos de revisión» constituyen una estrategia de creciente importancia para los clínicos frente a la eclosión de la literatura biomédica de hoy. En efecto la capacidad individual del profesional de leer y de absorber información es cada día más limitada por razones de accesibilidad a las numerosas revistas, de escasez de tiempo y de costo.

La reducción de grandes cantidades de información a síntesis seleccionadas y predigeridas puede ser útil para diversos profesionales de los servicios de salud. Sin lugar a dudas un buen artículo de revisión es una comodidad inapreciable para el atareado médico clínico.³⁵

Desafortunadamente los artículos de revisión publicados en revistas biomédicas son con demasiada frecuencia científicamente defectuosos, técnicamente ineficientes y —a veces— falsos. En efecto, rara vez se definen las estrategias para identificar y seleccionar la información primaria, por una parte, y la información recolectada se revisa a la ventura con muy poca o ninguna atención a una valoración sistemática de la calidad, por la otra. Parecería que los autores de los artículos de revisión no se han percatado aún de que la revisión de la literatura es una aventura científica y que el refinamiento adecuado de la información requiere de una exploración amplia y crítica, de un análisis y evaluación imparcial y de una síntesis inteligente.

A pesar de que la estrategia de revisar la literatura primaria se plantea como necesaria y oportuna, los lectores pueden despistarse si la revisión no cumple con estándares científicos. En consecuencia todo esfuerzo por establecer esas normas ayudará al clínico ávido de información a evaluar la calidad científica de la revisión.

Estas pautas de evaluación se concentran en cuatro aspectos fundamentales: a) Cuán clara es la definición de la pregunta clínica que origina la revisión (etiología, diagnóstico, tratamiento, pronóstico o prevención); b) Cuán exhaustiva es la estrategia de la revisión; c) Cuán apropiados son los métodos para seleccionar y valorar

los estudios primarios publicados; d) Cuán eficientes son los métodos para combinar los resultados y justificar las conclusiones.

El clínico debe aprender a leer en forma selectiva y en forma crítica las revisiones. Así como el uso de métodos defectuosos en un estudio original de diagnóstico o de tratamiento pueden invalidar los resultados, una revisión no científica de la literatura puede conducir a conclusiones incorrectas y peligrosas. Los autores de revisiones recolectan y analizan datos desde investigaciones primarias, lo cual se hace en forma subjetiva y subconsciente. La diferencia fundamental entre un estudio original (información primaria) y un estudio de revisión (información secundaria) está en el tipo de información y en la unidad de análisis y no en los principios científicos que se aplican.

La necesidad para el lector de contar con pautas de evaluación de las revisiones cobra máxima importancia cuando el clínico se enfrenta a distintas revisiones sobre un mismo tema las cuales llegan a conclusiones diferentes y hasta contradictorias. ¿A cual revisión creer en esas circunstancias?. La respuesta no es fácil, excepto que la «experiencia» del autor de la revisión y el «prestigio» de la revista que la publica no son suficientes para la credibilidad de la revisión. Numerosos estudios de la literatura médica muestran que la calidad científica de la mayoría de las revisiones publicadas, incluyendo aquellas en revistas de mucha respetabilidad, son pobres por cuanto no proveen argumentos suficientemente fuertes que apoyen claramente las conclusiones.

2. Pautas para evaluar la calidad de revisión

Una forma de enmarcar las pautas es utilizar una serie de preguntas que el lector le dirige al texto del artículo de revisión. Antes de entrar al detalle técnico de las preguntas es necesario aclarar que la intención de las pautas directrices es alentar el uso eficiente de la literatura médica mediante un escepticismo saludable.

A continuación se propone una serie de preguntas necesarias para una evaluación eficiente.

Pregunta N°1: ¿Se establecieron claramente las preguntas y los métodos de la revisión?

Pregunta N°2: ¿Se utilizaron métodos exhaustivos en la búsqueda de los artículos incluidos en la revisión?

Pregunta N°3: ¿Se utilizaron criterios explícitos para decidir qué artículos incluir y cuáles excluir de la revisión?

Pregunta N°4: ¿Se evaluó la validez de los artículos originales revisados?

Pregunta N°5: ¿Es reproducible la evaluación de los artículos originales revisados?
¿Está exenta de sesgos?

Pregunta N°6: ¿Se analizó la variabilidad de los hallazgos de los artículos originales revisados?

Pregunta N°7: ¿Se combinaron adecuadamente los hallazgos de los artículos originales revisados?

Pregunta N°8: ¿Se respaldaron las conclusiones de la revisión por los datos citados?

Este conjunto de ocho preguntas servirán poderosamente al clínico que busca respuesta —a través de la revisión— a algún problema de su práctica profesional relacionado con etiología, diagnóstico, pronóstico, tratamiento o prevención de enfermedades. La no respuesta o la respuesta insatisfactoria a las preguntas formuladas le servirá al clínico para descartar rápidamente artículos de revisión que son irrelevantes, científicamente poco sólidos, erróneos o falsos.

3. Estructura

Un artículo de revisión consta, en general, de nueve partes o secciones. Dentro de un esquema común la extensión de las secciones de un artículo de revisión podrán variar según el tema y el enfoque de la revisión. Cabe recordar que los artículos de revisión no llevan apéndice.

1. Título

Valen aquí los comentarios y normas definidos en el capítulo B.2.

2. Autor(es)

Valen aquí los comentarios y normas descritas en el capítulo B.2.

3. Resumen

Valen aquí los comentarios y normas descritas en el capítulo B.3. La estructura del resumen debe modificarse, adaptándolo a la estructura del capítulo C.4. al C.8 que se presentan a continuación.

4. Objetivo

Enfocar los factores más importantes considerados en la revisión; causa, diagnóstico, pronóstico, intervención terapéutica o prevención. Debe incluirse información sobre la población específica, la intervención o exposición y la prueba o desenlace (resultado) que hayan sido objeto de estudio.

5. Fuentes de datos

Resumir brevemente las fuentes y mencionar cualquier restricción. Las posibles fuentes abarcan expertos o instituciones que han investigado el tema, bases de datos

computarizados, índices publicados, registros, folletos, resúmenes, actas de conferencias, referencias extraídas de bibliografías en artículos o libros pertinentes, compañías o fabricantes de pruebas o sustancias.

En cuanto a bases de datos bibliográficos, es preciso indicar exactamente las palabras clave usadas en la búsqueda de información y cualquier restricción establecida (por ejemplo, idioma original, ámbito geográfico, período de tiempo que abarca, etc.).

6. Selección de estudios

Describir los criterios que rigieron la selección de estudios (criterio de inclusión y de exclusión) con detalles sobre la población particular, las intervenciones, los resultados o diseños metodológicos. Debe especificarse el método en que se basó la aplicación de los criterios: revisión a ciegas, revisión por consenso o por un grupo de revisores, etc..

7. Extracción de datos

Describir las pautas que se siguieron para recopilar los datos y para evaluar su calidad y validez (por ejemplo: criterios de inferencias causales). También hay que indicar el método empleado para seguir las pautas (por ejemplo: extracción independiente por varios observadores).

8. Síntesis de los datos

Presentar los resultados principales (cualitativos y cuantitativos) y esbozar los métodos con que se obtuvieron.

En los metaanálisis, deben indicarse los desenlaces o resultados más importantes que se conjuntaron así como la razones de productos cruzados (*odds ratios*), magnitudes de efectos y, si es posible, análisis de sensibilidad. Los resultados numéricos deben acompañarse de intervalos de confianza, siempre que proceda, y de los niveles exactos de significación estadística.³⁶

Las evaluaciones de pronósticos pueden incluir resúmenes de las características de supervivencia y de las variables relacionadas.³⁷

Deben indicarse las causas de variación entre estudios que se hayan identificado incluyendo, por ejemplo, las diferencias entre protocolos de tratamiento, coinervenciones, factores o variables de confusión, mediciones de desenlaces o resultados, duración del seguimiento y tasas de abandono.

9. Referencias

Valen aquí los comentarios y normas descritas en el capítulo B.13.. El número de referencias se recomienda en alrededor de cincuenta.

E. ESTILO EN LA COMUNICACIÓN CIENTÍFICA

M. A. Bobenrieth Astete

1. Redacción, gramática y estilo

Redactar, etimológicamente, significa compilar o poner en orden; en un sentido más preciso, consiste en expresar por escrito los pensamientos o conocimientos ordenados con anterioridad. Redactar bien es construir las frases con claridad, concisión, precisión, sencillez y naturalidad. Las ideas deben presentarse en alguna secuencia lógica.³⁸

La gramática es la ciencia que estudia los elementos de una lengua y sus combinaciones. El objeto de la gramática es el conocimiento sistemático y científico de la estructura del lenguaje según las normas consagradas por el uso. De esta manera, la gramática constituye un verdadero sistema de medios de expresión.³⁹

Para compensar la carencia de elementos extralingüísticos -mímica, gesticulación, contexto- al redactar un mensaje escrito surgen los signos de puntuación y de entonación, cuya misión es separar y clarificar las ideas.

La puntuación es el sistema de símbolos que ayudan al lector a comprender la relación estructural dentro de una frase u oración gramatical. Las marcas de puntuación en español pueden unir, separar, englobar, indicar omisiones, clasificar o terminar frases. La puntuación tiene una base más amplia que la decisión personal, que el capricho del escritor: esa base es la psicología.

Las marcas tienen estas funciones: el *punto* finiquita, la *coma* señala un intervalo, el *punto y coma* advierte una explicación inmediata, los *dos puntos* retardan para resaltar la idea que los sigue, los *puntos suspensivos* interrumpen enfáticamente, los *signos de admiración* muestran sobresalto, los de *interrogación* demandan aseveraciones, el *paréntesis* encierra alguna aclaración, el *guión* separa palabras formalmente o indica diálogo, las *comillas* rescatan la propiedad de los vocablos, la *diagonal* tiene diversos usos y la *diéresis* devuelve su sonido a la *u* muda.

De entre el conjunto de signos que sirven para puntuar se destacan cuatro: la coma, el punto y coma, el punto y el punto final.⁴⁰

La *coma* es el signo ortográfico que sirve para indicar la división de las frases o miembros más cortos de la oración o del período y que también se emplea en aritmética para separar los enteros de las fracciones decimales. Es pues una de las pausas que pueden hacerse al hablar y por lo común la más corta de ellas.

El *punto y coma* sirve para separar oraciones entre cuyo sentido hay proximidad y, por excepción, frases largas semejantes en serie. El punto y coma sirve para separar períodos relacionados entre sí, pero no enlazados por una preposición o conjunción.

El *punto* es el signo ortográfico con que se indica el final del sentido gramatical y lógico de un período o de una sola oración. Se pone también después de toda abreviatura. El punto separa oraciones cuando los pensamientos que ellos contienen, aunque relacionados entre sí, no lo están de modo inmediato.

Cuando el pensamiento se ha desarrollado en una o más oraciones, formando un todo que se llama «párrafo», se pone punto final. Es decir, cuando lo que se ha expresado tiene sentido completo.

La *ortografía* es la parte de la gramática que da reglas para el uso correcto de las diferentes letras, para el empleo del acento ortográfico y para la debida colocación de los signos de puntuación.⁴¹

La *estilística*, complemento de la gramática, añade a la corrección en el uso del lenguaje, la precisión, la elegancia, la claridad y la armonía.

La palabra estilo proviene del latín «stilus», término con el cual griegos y romanos denominaban el punzón de que se valían para escribir en tablas enceradas. De aquí pasó a designar el conjunto de rasgos que caracterizan la obra de un autor, escuela, género artístico, época o período, diferenciándolo de los demás.

El estilo es el orden y el movimiento a que se somete lo que pensamos. Es el modo característico de hacer, de escribir; es el sello del espíritu del autor sobre las formas. Pero más allá de lo dicho, el buen estilo ha de reunir una serie de cualidades, tales como la claridad, la concisión, la sencillez, la naturalidad, la objetividad y la originalidad.

Ernest Renán, filólogo, filósofo e historiador francés escribía ya en el siglo pasado: «La regla del buen estilo científico es la claridad, la perfecta adaptación al asunto, el completo olvido de sí mismo, la abnegación absoluta. Es también la regla para escribir bien sobre cualquier material».

El estilo reconoce factores subjetivos procedentes de lo peculiar y privativo de un autor, que es como el sello de su personalidad artística. Este hecho hizo decir a George Louis Leclerc, Conde de Buffon, en su discurso de recepción en la Academia francesa en 1753: «El estilo es el hombre mismo». No hay sólo estilos individuales, sino también de escuela, de generación y de época.

Habida consideración de todo lo anterior, el estilo debe atraer e interesar al lector, de manera tal que siga con atención el hilo del relato. El estilo árido, pesado, exagerado, ambiguo induce a abandonar precozmente la lectura del artículo. El genio de François Marie Arouet —mejor conocido como Voltaire— escritor, poeta y filósofo francés del siglo XVIII nos legó su frase: «Todos los estilos son buenos, excepto el aburrido».

Junto a estos factores subjetivos, procedentes del autor, de su orientación y de su época, hay exigencias objetivas que dependen del carácter y del objetivo de la obra. Cada tipo de creación escrita responde a un objetivo especial, lo cual impone ciertas trabas a la libre exteriorización del estilo personal. Dentro del estilo literario no es exactamente igual el estilo de una poesía, de un drama de una novela o de un artículo periodístico. Mayor aún es la diferencia entre el estilo literario (poesía, drama, novela, historia, biografía, periodismo y oratoria) y el estilo científico, tanto en sus aspectos formales como en cuanto al fondo.⁴¹

2. Diferencias entre el estilo científico y el estilo literario *

En la comunicación escrita, al igual que en la hablada, la intención o las consecuencias deseadas determinan el estilo o forma de expresión. La comunicación es un proceso (del latín *processus*, participio pasado del verbo *procedere* «avanzar») y, por lo tanto, algo continuo y cambiante, cuyas partes interactúan e influyen una sobre otra.⁴⁰

Mediante la escritura se pretende «fijar» o «congelar» ese proceso fluido y darle estabilidad. Sin embargo, en el proceso participa el receptor del mensaje, quien a su vez lo interpreta de acuerdo con una serie individual, única e infinita de factores tanto objetivos como subjetivos: de entre ellos, su conocimiento del lenguaje y del tema, contexto cultural, estado de ánimo, prejuicios, predisposiciones.

De ello se desprende que el escritor científico tiene que hacer todo lo posible por evitar ambigüedades y giros que den lugar a interpretaciones erróneas. Para ceñirse a un «estilo científico» de alcance tanto nacional como internacional, deben tenerse en cuenta las siguientes diferencias entre éste y el estilo literario:

* Tomado parcialmente de un documento inédito de la Dra. Eglá Blauin, editora jefe del Boletín de la OSP. OPS/OMS, Washington, D.C. 1991

*Estilo científico**Estilo literario*

- | | |
|---|---|
| 1. Primariamente informativo y persuasivo. | - Primariamente entretenido y recreativo. |
| 2. Racional, factual. | - Emotivo. |
| 3. Objetivo. | - Subjetivo. |
| 4. Impersonal, desinteresado. | - Personal, interesado. |
| 5. Sacrifica el deseo de originalidad estilística a la expresión concisa y clara de los datos. No obstante, el efecto final debe ser agradable. | - Permite originalidad de estilo: ironía, choque, anécdota, humor, exageración, ovación, entusiasmo. |
| 6. Trata de comunicar lo concreto. | - Trata de comunicar lo abstracto y lo infame. |
| 7. Descubre una realidad vivida. | - Crea una realidad no vivida. |
| 8. Presenta información comprobable | - Inventa hechos y situaciones. |
| 9. Muestra. Ocasionalmente demuestra | - Describe, sugiere, insinúa. |
| 10. Está libre de jerga y pomposidad. | - Usa jerga para crear ambientes, acentos y personajes. |
| 11. Evita modismos lugareños. Si ha de usarlos, los identifica a pie de página. | - Se vale de idiotismos y de exoticismos. |
| 12. Es transparente, no oscurece la idea; elimina toda ambigüedad. | - Puede oscurecer el mensaje para crear ecos subconscientes, poéticos; para convertir al lector en «cocreador». |
| 13. Va directamente al punto y especifica. | - Permite rodeos, descripciones extensas, exclamaciones, inespecificidad. |
| 14. Presenta un punto de vista sin transparentar predilecciones personales o emociones. | - Permite cambiar el punto de vista según los personajes o emociones del autor. |
| 15. Distingue hechos (datos) de opiniones. | - Puede mezclar hechos con opiniones. |
| 16. Invita al diálogo y a la discusión | - Invita a la introspección y al goce interior de imágenes, ideas y sensaciones. |
| 17. Usa términos y formas convencionales | - Prefiere lo arbitrario. |

-
- | | |
|---|---|
| 18. Sigue un orden preestablecido. | - No tiene que seguir un orden especial. |
| 19. Obedece al concepto lineal de tiempo y espacio. | - Puede violar ese concepto. |
| 20. Explica los conceptos difíciles. | - Deja la explicación a cargo del lector. |
| 21. Enseña, adoctrina, adiestra. | - Deleita, place el ánimo y los sentidos. |
-

3. Cualidades del buen estilo

El buen estilo científico posee ciertas cualidades esenciales las cuales se describen a continuación.

- a. **CLARIDAD.** Se es claro cuando el escrito penetra sin esfuerzo en la mente del lector, cuando se expresa con lisura, sin rebozo ni simulación. La claridad consiste en que la idea se exponga de manera que evite interpretaciones erróneas y solo dé a entender lo que el autor quiere decir.⁴³ José Ortega y Gasset, filósofo y escritor español de la Generación Novecentista, escribía: «La claridad es la cortesía del científico».

Para lograr claridad en el estilo no basta con tener las ideas claras. Es necesario que la construcción de la frase y del párrafo responda al orden lógico-psicológico de las ideas. Para asegurar esto último conviene ligar las ideas entre dos o más frases; además tales ideas deben presentarse en una secuencia, según su importancia.

Para conseguir la debida cohesión en un párrafo o período, debe procurarse ligar la idea inicial de una frase a la idea final de la frase precedente o a la idea general -dominante- de dicho párrafo.

El pronombre relativo debe colocarse cerca de su antecedente.

Debe evitarse el uso de palabras rebuscadas, afectadas.

A diferencia del estilo literario, donde las palabras son muchas veces protagonistas del verso o de la prosa, los vocablos en el estilo científico han de ser simples vehículos del pensamiento, al servicio del mismo; las palabras debe ser «transparentes».

El oscurantismo literario y el confusionismo expresivo no tienen cabida en el estilo científico. El estilo científico es por esencia utilitario. El escritor científic-

co es un hombre que se sirve de las palabras. El estilo científico es la expresión de una actitud del espíritu del investigador: hay claridad cuando la palabra pasa a través de nuestra mirada como el sol a través del vidrio limpio de la ventana.

- b. **CONCISIÓN.** Se es conciso cuando se usan sólo las palabras indispensables, precisas y significativas para expresar lo que se quiere decir.⁴³

Concisión implica brevedad, centrando el mensaje en lo esencial, en lo que el autor considera verdaderamente importante. Conciso no quiere decir lacónico, sino denso. Estilo conciso es aquel en que cada frase, cada línea o cada palabra están preñadas de sentido. Lo contrario es la vaguedad, la imprecisión, el exceso de palabras; lo que se llama «retórica» en su significado de artificio excesivo, de rebuscamiento en el lenguaje.

William Shakespeare, insigne poeta inglés y uno de los primeros dramaturgos del mundo, escribía a comienzos del siglo XVII: «La brevedad es el alma del ingenio». A su vez Anton Chejov, novelista y dramaturgo ruso, declaraba tres siglos más tarde: «La brevedad es hermana del talento». En España Pío Baroja, novelista de la Generación del 98, escribía en *La Institución y el Estilo. Memorias*: «Con el tiempo cuando los escritores tengan una idea psicológica del estilo y no un concepto burdo y gramatical, comprenderán que el escritor que con menos palabras pueda dar una sensación más exacta y completa, es el mejor».

La falta de concisión es un defecto general de los que empiezan a escribir. Es preciso limpiar el estilo, quitar la paja del trigo, podarlo hasta que no se pueda decir más concisamente lo que hayamos dicho. Es preciso evitar lo superfluo, la verborrea, el añadido de ideas secundarias que no agregan nada a la idea matriz, sino que más bien la debilitan.

A. Albalat, estilista francés, declaraba en *L'art d'écrire* «No se olvide que las frases están hechas las unas para las otras y que es su encadenamiento apretado el que origina una de las bellezas generales del estilo. Que las frases no aparezcan injertadas, sino engendradas; no ficticiamente yuxtapuestas, sino lógicamente deducidas».

- c. **PRECISIÓN.** Se es preciso cuando se usa un lenguaje directo, sin términos ambiguos ni expresiones confusas o equívocas. La precisión apunta a que la expresión científica sea rigurosamente exacta.

Para que el estilo sea preciso se parte del hecho de que si bien es cierto que hay sinónimos, no es menos cierto que hay una palabra específica para un objeto, una situación y un lugar también específico. El autor debe buscar siempre la palabra más adecuada.

Una forma de mejorar la precisión en el lenguaje consiste en evitar el abuso de las «palabras fáciles». Éstas son vocablos de muy amplia significación, voces vagas, imprecisas, incoloras que, a fuerza de servir para todo, terminan por servir poco o por no servir apenas para nada. Entre tales «palabras fáciles» están los pronombres, los adverbios, las conjunciones y algunos verbos. Además están otros vocablos indeterminados a manera de «comodines verbales» que acuden a la pluma del autor. De entre este último grupo se destacan las palabras «cosa», «algo», «esto» y «eso»; el sentido de éstas suele ser vago, impreciso, hasta vulgar y trivial.

También conspira contra la precisión el abuso de la voz pasiva; del pronombre relativo «que»; de los verbos «ser», «estar», «encontrarse», «haber» (impersonal) y «tener»; también caen aquí los verbos «hacer», «poner», «decir» y «ver». Tampoco debemos olvidar el abuso del verbo «producir» y de las palabras «tema», «asunto», «cuestión», «tópico».

- d. **SENCILLEZ Y NATURALIDAD.** Se es sencillo cuando empleamos lenguaje de uso común, sin caer en la vulgaridad. Se es natural cuando usamos lenguaje y expresiones propias de nuestro acervo personal.

Sencillez es huir de lo enrevesado, de lo artificioso, de lo complicado, de lo «barroco»; sencillo es el autor que utiliza palabras y frases de fácil comprensión. La sencillez -se ha dicho- es el sello de la verdad. Natural es el autor quien, al escribir, utiliza el vocabulario de su habitual modo expresivo.

Lo contrario de lo sencillo es lo artificioso; lo contrario de lo natural es lo afectado, lo rebuscado. Vienen aquí a nuestra mente las palabras de Friedrich Nietzsche, célebre filósofo alemán de fines del siglo XIX: «La sencillez y naturalidad son el supremo y último fin de la cultura».

4. Papel de la lógica y de la gramática en la redacción correcta

Para redactar en forma organizada, clara, precisa y sencilla se requieren dos factores fundamentales:

- a. Orden en las ideas, lo cual se apoya en la *lógica*; ésta consiste en la disciplina que estudia los principios formales del conocimiento humano. La lógica tiene que ver con las doctrinas del concepto, del juicio, del silogismo y del método.⁴⁴ En consecuencia nos ayuda a ordenar las ideas para exponerlas con claridad.

El conocimiento se adquiere y se transmite a través de sensaciones, percepciones, representaciones, conceptos, juicios y razonamientos. Estos tres últimos elementos son formas más complejas del pensamiento.⁴⁴

El concepto es la representación mental del objeto, el cual puede definirse, dividirse y clasificarse. El juicio es la afirmación o negación de algo, que surge al establecer una relación entre dos conceptos. El razonamiento es una forma de obtener o demostrar un conocimiento, a partir de verdades preestablecidas.

- b. Utilización de las reglas básicas de la comunicación escrita lo cual se apoya en la *gramática*, es decir aquella parte de la lingüística que estudia las clases de palabras, su composición y accidentes, y su función en la oración.

La lengua es el sistema más eficaz de comunicación humana que se caracteriza porque emplea signos, compuestos por dos partes inseparables: el *significado*, que es el contenido, la idea, y el *significante*, que es la forma, ya sea sonora ó gráfica.

La palabra actúa como unidad mínima significativa de la lengua. La palabra se compone de unidades más pequeñas que son de dos tipos: a) la raíz que es la parte invariable o «lexema» de la palabra; b) la terminación que es la parte flexible o «gramema» de la palabra.

Los accidentes gramaticales más importantes son: a) el *genero* (grupo al que pertenecen los seres); b) el *número* (señala si son uno o más de los seres que hablan o participan en la acción); c) la *persona* (ayuda a precisar quién habla, con quién se habla, de quién se habla).

De la lengua pueden distinguirse tres usos: a) el científico, en que interesa más el contenido que la forma; b) el cotidiano, que busca resultados prácticos; c) el literario, que busca recrear el universo vivencial del lector.

La lectura crítica intenta la comprensión del mensaje del autor y la aceptación o rechazo de las ideas vertidas en el texto a través de la identificación de los signos gráficos del lenguaje.

5. Cincuenta y un consejos útiles para evitar vicios de mal estilo

El redactar bien es un asunto no solamente de talento sino también de destrezas. La inteligencia, la cultura y la aptitud juegan un papel fundamental para representar adecuadamente las ideas en palabras y otros signos; la habilidad y el arte basados en sabiduría, en práctica y en experiencia constituyen el complemento necesario para lograr un producto de comunicación escrita pertinente y atractivo, de alcance universal sin perder el tono personal.⁴⁵

Miguel de Unamuno, filósofo, novelista, poeta, dramaturgo, ensayista y filólogo de la Generación del 98, dejó huella de su pensamiento, el cual parece oportuno recor-

dar: «Sucede con frecuencia que llega uno a encontrar una idea que buscaba, llega a darle forma, es decir a obtenerla, sacándola de la nebulosa de percepciones oscuras a que representa, gracias a los esfuerzos que hace para presentarla a los demás».

La conjunción de talento y de destrezas transforma el escribir en una experiencia alegre y gratificadora. Un signo inequívoco de esta transformación es cuando el escribir depende menos del talento y del humor para llegar a convertirse en un hábito productivo y agradable.⁴⁶

Más allá del talento y de las destrezas, el escribir está lleno de escollos ocultos y de peligros latentes los cuales asoman a la superficie con demasiada frecuencia bajo la forma de errores, de yerros y de faltas, las cuales le restan claridad y comprensión al mensaje aunque no siempre es posible establecer marcos bien delineados e inequívocos sobre los diversos tipos de faltas.⁴⁷

Nos anima a escribir este capítulo la experiencia ganada mediante la lectura crítica de unos tres mil manuscritos sometidos para publicación como artículos científicos originales en el *Boletín de la OSP* a lo largo de 13 años en la Región de las Américas de la OMS. De este universo, no más de un 25% llegó a la meta de la publicación, después de un laborioso proceso de crítica, de producción de textos revisados y corrección de estilo. A esto se suma la experiencia docente de más de 30 cursos intensivos, talleres y seminarios sobre actualización en escritura científica en España, Chile, Perú, Venezuela, República Dominicana, Cuba, Nicaragua, Costa Rica, Honduras, Guatemala, México y los EE.UU. donde se sometieron a «diseción» -tanto de contenido como de estilo- unos doscientos artículos científicos originales y de revisión seleccionados al azar, de las principales revistas biomédicas de España y de los países del Hemisferio Occidental.

A lo largo de estos años, varios alumnos y autores nos alentaron a escribir y publicar nuestra experiencia sobre estas materias; lo hemos hecho con la esperanza de que sea de real utilidad para el lector interesado.

Las cincuenta y una reflexiones que siguen constituyen verdaderas «alertas» útiles, consejos que se dan para redactar mejor. No intentan -de manera alguna- ahorrar al lector el proceso de «ensayo y de error» indispensable en la formación de un escritor. Sólo proponen algunas pautas directrices para acortar este proceso y transformar los errores en verdaderas experiencias de aprendizaje y de desarrollo.⁴⁸ Intentan, además, confirmar nuestra convicción de que lo que se lee y se comprende sin mayor esfuerzo, se ha escrito siempre con un gran esfuerzo.

Cuán hondo y oportuno calan los versos de Antonio Machado, el poeta de la generación del 98: "Caminante, son tus huellas / el camino, y nada más; / caminante,

no hay camino, / se hace camino al andar". Miguel de Cervantes, príncipe de los ingenios españoles, escribía «No se toman truchas a bragas enjutas». Por su parte George Clemenceau, célebre médico, político y periodista francés, escribía al respecto en los años veinte: «La vida es una perseverancia».

1. Aclare sus ideas antes de escribir el texto; una vez que vea el tema con claridad, hágase un esquema que le sirva de guía para la escritura de frases, de párrafos y de capítulos. Viene a nuestra mente aquella afirmación de Oscar Wilde, poeta y dramaturgo inglés del siglo XIX: «No existen más que dos reglas para escribir: tener algo que decir y decirlo bien».
2. Recuerde que su texto requiere de un orden en las ideas, el cual se apoya en la lógica. Además debe utilizar las reglas básicas de la comunicación escrita, la cual se apoya en la gramática.
3. Organice jerárquicamente sus ideas y preséntelas en alguna secuencia lógica que facilite su comprensión. Destaque la idea principal en cada capítulo del texto y luego desarrolle las ideas secundarias; exponga claramente la idea que ha escogido como fundamento y describa luego las ideas modificantes.
4. Anote las ideas y plántelas en un orden lógico deseable que sirva al objetivo del escrito. La idea básica debe destacarse aun cuando no vaya al principio de la frase.
5. Asegúrese de que cada capítulo o sección del texto tenga un planteamiento, un desarrollo y un cierre.
6. Evite a toda costa las ideas superfluas e irrelevantes. En consecuencia, tache lo que no está relacionado con la idea fundamental de la frase o párrafo.
7. Adapte su mensaje a la experiencia de un lector promedio; utilice términos que el lector pueda comprender.
8. Defina los términos claves del estudio; el estilo de la definición debe ser afirmativo, preciso, claro y sencillo. Aquí tienen plena vigencia las palabras de José Ortega y Gasset, filósofo y escritor español de la Generación Novecentista: «Una lengua es un sistema de signos verbales merced al cual los individuos pueden entenderse sin previo acuerdo, al paso que una terminología sólo es inteligible si previamente el que escribe o el que habla y el que lee o escucha se han puesto individual o colectivamente de acuerdo sobre el significado de los signos».
9. Las palabras son la «materia prima» del artista, las «herramientas» del escritor. En consecuencia, un pilar fundamental del autor de un escrito es su conocimiento del vocabulario. El empleo de la palabra exacta, apropiada y adecuada es una de las reglas fundamentales del estilo.

10. Utilice palabras unívocas; nunca emplee un término si no está seguro de lo que verdaderamente significa y de cómo se escribe. La terminología es el conjunto de términos que se utilizan en un campo determinado del saber, un campo especializado, en forma distinta a como se usan en el lenguaje común y que se han fijado por uso o por normalización.
11. No ponga más palabras que las indispensables; evite palabras innecesarias, las cuales conducen a vicios comunes en la redacción biomédica, tales como la grandilocuencia, la verborrea y la redundancia.
12. Tenga presente que las dos palabras que comandan el idioma son el verbo y el sustantivo. Las restantes partes de la oración son complementarias y no se debe abusar de ellas.
13. Tenga siempre a mano tres buenos diccionarios de la lengua española: a) uno etimológico; b) otro de sinónimos y contrarios; c) un tercero de dudas del idioma; consúltelos ante cualquier dificultad acerca del significado y la ortografía de las palabras.
14. El empleo del adjetivo, entendiéndolo por tal aquella parte de la oración que se junta al sustantivo para calificarlo o determinarlo, debe ser lo más exacto posible. Evite, pues, la duplicidad innecesaria de adjetivos.
15. No dé proporciones excesivas a las cosas llevándolas más allá de los límites de lo natural, de lo razonable, de lo justo, de lo conveniente y de lo verdadero. Los hechos narrados limpiamente convencen más que las exageraciones.
16. Ubique los adverbios -sean éstos de tiempo, de lugar, de modo, de cantidad, de afirmación, de negación o de duda- cerca del verbo al cual se refieren; la exposición resultará así mas clara.
17. No abuse de los pronombres. Sobre todo tenga sumo cuidado con el empleo del pronombre posesivo «su» -apócope de suyo, suya, suyas, suyos- el cual suele producir ambigüedad ya que por su apócope pierde el género y es causa frecuente de anfibología (doble sentido).
18. Tenga muy en cuenta que «la puntuación es la respiración de la frase». No olvide que una causa frecuente de falta de claridad es la mala puntuación.
19. Esté alerta con el uso de barbarismos y de solecismos.⁴⁸ Se llama barbarismo a un vicio de dicción que consiste en escribir o pronunciar mal las palabras, o en emplear vocablos impropios. Solecismo es falta a las reglas dadas acerca del oficio y uso de las distintas partes de la oración. Mientras que el barbarismo atañe a las otras partes de la gramática, el solecismo se refiere a la sintaxis.

20. Evite el uso de extranjerismos, es decir voces, frases o giros de un idioma, extranjero empleado en español. Aquí se incluyen en forma mayoritaria los anglicismos, los americanismos y los galicismos; en proporción menor los germanismos, latinismos, grecismos, italianismos, portuguesismos, galleguismos y catalanismos.
21. Prefiera la expresión simple a la compleja evitando el uso de vocablos rebuscados con exceso de atildamientos que degeneran en afectación del lenguaje y del estilo.⁴⁹
22. Prefiera el vocablo culto frente al vocablo vulgar. Los vulgarismos son una forma de barbarismo que consiste en el empleo de palabras o expresiones propias de la lengua hablada por el estrato menos culto de la población, inadmisibles en la lengua profesional usual. Los vulgarismos no son sino los barbarismos más evidentes utilizados por personas que carecen, por lo general, de cultura y de formación gramatical. Sin embargo, en este vicio no cae sólo la gente vulgar sino muchas personas que presumen de cierta cultura.
23. Use formas verbales activas. Recuerde que el idioma español tiene preferencia por la voz activa. La voz pasiva suele imponerse por desconocimiento del agente activo, porque hay cierto interés en ocultarlo, porque le es indiferente al escritor y por la influencia del idioma inglés.⁵⁰
24. No abuse de los incisos y de los paréntesis. Ajústelos y procure que no sean excesivamente extensos.
25. Evite las repeticiones excesivas y malsonantes. Repetir puede ser legítimo solamente cuando se quiere fijar la atención sobre una idea y siempre que no suene mal.
26. Amplíe su vocabulario. Enriquezca su escrito mediante el uso de vocablos o expresiones que tienen una misma o muy parecida significación. Si emplea sinónimos, procure que no sean muy raros, de uso poco común o poco frecuente.
27. Manténgase alerta contra la sintaxis defectuosa, un vicio de redacción muy común. Al construir las frases tenga en cuenta el arreglo sintáctico: sujeto, verbo y complemento en un orden lógico.
28. Evite tecnicismos inoportunos y alardes vanos de erudición. Aclare el significado de las voces técnicas cuando no sean de uso común.
29. No ubique el verbo al final de la frase. Esta situación es frecuente en la lengua alemana pero no encaja bien en la gramática española.
30. Procure ligar la idea inicial de una frase a la idea final de la frase anterior para facilitar la debida cohesión entre las oraciones.

31. La construcción armoniosa exige evitar las repeticiones malsonantes (cacofonía), al igual que la uniformidad e igualdad de tono como resultado de una pobreza de vocabulario (monotonía) y las asonancias.
32. Evite las transiciones bruscas entre distintos párrafos. Procure «fundir» las ideas y conceptos con habilidad para que las transiciones sean suaves y efectivas.
33. Tenga presente que el estilo directo tiene mas fuerza y causa mas impacto que el estilo indirecto.
34. Escriba de preferencia con frases cortas. La psicología enseña que una frase compuesta de 20 palabras logra hasta un 90% de comprensión en los lectores; a medida que aumentan las palabras en la misma frase, la comprensión del lector disminuye. Sin embargo evite la sucesión monótona de frases cortas ininterrumpidas lo cual incluye el abuso de «punto y seguido»; igualmente evite los períodos extensos los cuales resultan muchas veces ampulosos y vagos.
35. Recuerde que el lenguaje es el medio de comunicación por excelencia y que las cualidades fundamentales del estilo son: la claridad, la concisión, la precisión, la sencillez y la naturalidad.
36. Haga que su estilo sea atractivo para que capte la voluntad del lector. Utilice la variedad dentro de la unidad.
37. Evite que su estilo sea superficial, ordinario, lacónico y tremendista.
38. Además de cuidar el estilo, tenga en cuenta el tono, que es el estilo adaptado al tema y a los resultados de la investigación; el «tono» o «entonación» es la adecuación de la forma al fondo; el tono acomoda las palabras y las frases a los hechos y las opiniones que se expresan por escrito.
39. Manténgase aparte de las «frases hechas» y de los «lugares comunes» lo cual se manifiesta en expresiones vulgares o triviales.
40. Sea muy cuidadoso con el uso de metáforas, tropo que consiste en trasladar el sentido recto de las voces en otro sentido figurado en virtud de una comparación tácita. La metáfora solo vale cuando añade fuerza expresiva y precisión a lo que se escribe; en caso contrario la metáfora suele confundir al lector.
41. Recuerde que una cualidad esencial de la escritura científica es la precisión. Huya, en consecuencia, de la sugestión sonora de las palabras.
42. Establezca en su estilo un balance adecuado entre imaginación y objetividad; entre entusiasmo y reflexión; entre seguridad y escepticismo; entre gallardía intelectual y humildad.

43. Escriba para expresar, no para impresionar. Ud. es responsable de darse a entender bien por medio de la palabra escrita; no cabe, en consecuencia, el encajar una idea u opinión en la mente del lector por medio de persuasión indebida o de manera conmovedora.
44. Tenga en cuenta que el estilo científico no es un ornamento que transforma lo simple en grandioso; más bien, su función es la opuesta y consiste en volver comprensible lo complejo. En última instancia es el estilo lo que da significado a los pensamientos. Si, como escribió Albert Einstein, «la ciencia no es más que el pensamiento cotidiano refinado», el estilo nos sirve para lograr dicho refinamiento.
45. Prevenga el riesgo de padecer el «Síndrome de la Celebridad Intelectual» descrito por J.T. Winkler,⁵¹ cuyos síntomas principales son: a) Personalización de las ideas con una alta densidad de pronombres personales («yo», «me», «mi», «nosotros» etc.); b) Argumentación mediante anécdotas favorables seleccionadas; c) Exageración del significado; d) Lenguaje portentoso y pretencioso; e) Publicación prematura.
46. Su estilo es perfectible. Una forma de lograrlo es leer asiduamente con atención y reflexión a los buenos escritores. En efecto la lectura atenta de los maestros de la literatura, tanto de la clásica como de la moderna y de la contemporánea, ayuda eficazmente a mejorar la capacidad de escribir.
47. Una vez terminado el texto del artículo délo a leer a dos personas de su confianza que entiendan del tema para su lectura crítica. Este ejercicio le permitirá corregir las insuficiencias, los excesos y lo ininteligible.
48. Una vez producido el texto corregido, pida a alguien con formación lingüística que lo revise para comprobar la sintaxis, la puntuación y la ortografía.
49. Resuelva la duda de por qué escribir, qué escribir, cuánto escribir, cómo escribirlo y dónde publicarlo ubicándose en el interés y en el lugar del lector inteligente e interesado.
50. Al poner por escrito los resultados de la investigación y discutir su significado, tenga presentes las tres grandes amenazas a su validez y veracidad: a) El error, a veces involuntario e inevitable; b) El sesgo como error sistemático y generalmente detectable; c) El fraude como engaño deliberado del autor al lector confiado.
51. Si realmente no tiene nada importante que comunicar, contenga su ímpetu. George Eliot, novelista del realismo inglés, escribió a fines del siglo pasado: «Bendito es el hombre quien no teniendo nada que decir se abstiene de evidenciar ese hecho en palabras».

F. LECTURA CRITICA DE ARTÍCULOS CIENTÍFICOS

M.A. Bobenrieth Astete

1. El clínico como juez crítico de artículos

El crecimiento exponencial de la literatura biomédica en los últimos cincuenta años mencionado en la introducción de esta obra —especialmente en la forma de artículos originales— ha tenido gran influencia en la práctica clínica. En efecto la publicación de resultados de investigaciones sobre etiología, diagnóstico, tratamiento, factores pronósticos y prevención de enfermedades en revistas alcanza rápidamente la comunidad médica nacional. Una vez incorporada esa información en bases de datos especializados, está al alcance de la comunidad médica internacional.

Un artículo publicado en una revista biomédica supone para el lector incauto que esa investigación es sólida y que vale clínicamente. Lamentablemente el solo nombre del autor y el supuesto prestigio de la revista no constituyen garantía alguna de validez científica de la publicación. Los estándares de investigación dejan mucho que desear desde el punto de vista del diseño, del análisis de la información, de la estructura y de la redacción del informe de la misma. En efecto muchas de las revistas biomédicas no disponen de un verdadero sistema de arbitraje por expertos, ni de consejos editoriales idóneos y ágiles. Aquellas revistas que anuncian disponer de estos mecanismos de control de calidad del contenido tampoco garantizan necesariamente la validez del diseño o de las conclusiones. De hecho el funcionamiento de los consejos editoriales suele ser irregular con miembros que cooperan más por amistad y buena voluntad que por verdadera capacitación metodológica. Los abundantes ejemplos de artículos originales publicados en casi todos los números de revistas sin estándares de calidad básicos de diseño, con análisis parciales e incorrectos de los datos y plagados de defectos severos de redacción hablan por sí solos de un gran problema no resuelto de desinformación.

Nunca se insistirá suficientemente sobre la importancia de un diseño y un análisis acertados. Claramente las conclusiones de un estudio dependen de la corrección de los métodos. Si las conclusiones no son confiables debido a metodologías defec-

tuosas, entonces el estudio no puede tener valor clínico. Peor aún, puede ser clínicamente dañino a causa de una práctica profesional basada en conclusiones falsas y engañosas; un estudio clínicamente dañino no es ético. La aceptación profesional de los hallazgos dudosos tiene entre sus causas una ignorancia de aspectos metodológicos fundamentales.

Las consideraciones precedentes nos permiten concluir que hoy en día el clínico necesita conocimientos y destrezas básicas para evaluar en forma inteligente y crítica los trabajos que se presentan a reuniones científicas y aquellos otros que se publican como artículos originales. Esto le permitirá detectar errores y sesgos de diseño, de ejecución, de análisis, de presentación, de interpretación y de omisión. Diversas revisiones de la literatura informan que alrededor de la mitad de los artículos examinados contienen errores estadísticos de los tipos antes mencionados, en diverso grado, con las consiguientes implicaciones clínicas y éticas. No cabe duda de que los estándares de metodologías de investigación, y de estadística en especial, podrían mejorarse grandemente si las revistas biomédicas cesaran de publicar artículos metodológicamente deficientes; sin embargo no se percibe un movimiento claro en el sentido de ese mejoramiento.

En consecuencia el médico clínico debe ser cauteloso en aceptar lo publicado como válido o como cierto y asumir de una vez su papel de juez crítico de lo publicado. Los diversos capítulos de este libro apuntan a desarrollar directamente esa capacidad de lectura crítica. Solamente queda pendiente el disponer de algunas pautas, de guías concretas para efectuar esa evaluación en forma metódica. La sección que sigue intenta cumplir precisamente con ese objetivo.

2. Listas de revisión

Un examen de la literatura sobre este tema muestra algunos libros sobre investigación en ciencias de la salud^{17, 54}; sobre estadística en investigación biomédica⁵⁵, sobre epidemiología clínica²⁹, sobre investigación en enfermería⁵³ y sobre metodologías de investigación⁵⁶, los cuales incluyen capítulos especiales sobre evaluación crítica de la investigación publicada. El objetivo general de estos capítulos es proporcionar listas de elementos que se deben revisar en cada investigación publicada. Estas listas se organizan agrupando frases en tono afirmativo o en forma de preguntas sobre las etapas del proceso de investigación.

En algunos casos se agregan elementos relativos a las partes de la estructura del artículo original y de su estilo; esto es de suma utilidad ya que, en la práctica, contenido y forma están indisolublemente unidos en el artículo original.

Algunas revistas biomédicas utilizan el sistema de arbitraje por expertos, llamado «peer review system» en idioma inglés⁵⁷. Este consiste en un sistema de evaluación por especialistas en las diversas materias, los cuales examinan el valor científico del manuscrito y la utilidad de su publicación.⁵⁸ A fin de ayudar a los árbitros a cristalizar sus puntos de vista y a tomar decisiones, las revistas más serias suelen emplear listas comprobatorias y verdaderos formularios diseñados especialmente para el objetivo descrito. El uso de formularios favorece la homogeneización de la evaluación y disminuye la tasa de desacuerdos entre árbitros.⁵⁹ Sin embargo no pueden reemplazar la experiencia, el sentido común ni el examen detallado del árbitro.⁶⁰

La pauta N° 1 del Apéndice V se presenta como una descripción sucinta del formulario del árbitro que utiliza el Boletín de la Oficina Sanitaria Panamericana, revista mensual sobre temas de salud pública internacional de la Organización Panamericana de la Salud. Luego se describe brevemente el contenido de 5 capítulos de otros tantos libros, sobre evaluación crítica de la investigación publicada, cada uno de los cuales incluye una lista de revisión.

El hecho de que exista más de un formulario de revisión indica que ninguno en particular puede satisfacer exactamente las necesidades y requisitos de las revisiones. Esto es comprensible ya que la investigación en salud y la comunicación escrita de la misma presenta una gran diversidad, consecuencia natural de la variedad de temas, de énfasis, de disciplinas, de problemas, de enfoques, de diseños, y de metodologías.

El Apéndice V incluye además una reproducción de las listas de revisión de seis fuentes seleccionadas. En el primer y segundo caso se transcribe un facsímil del texto completo; en los otros se limita a incluir la lista ordenada y abreviada de items con ajustes editoriales menores tendientes a su mejor comprensión. De cualquier manera el lector interesado deberá consultar la fuente original que se cita en la sección de «Referencias» de esta obra.

3. Sinopsis de seis pautas de evaluación

PAUTA N° 1

Formulario del Arbitro. Boletín de la Oficina Sanitaria Panamericana.

Origen y Objetivos. La necesidad de facilitar el trabajo de 320 profesionales especialistas que actuaban en 1992 como miembros del sistema de revisión por árbitros de los manuscritos sometidos para publicación como artículos originales. Estos cubrían

alrededor de 72 disciplinas del saber y áreas de problemas del vasto campo de la salud humana individual y colectiva.

Este formulario del árbitro es un «ayuda memoria» para los revisores quienes, siendo autoridades reconocidas nacional e internacionalmente en el quehacer de su especialidad médica, no siempre tienen un dominio metodológico del complejo proceso de investigación científica dentro del marco de política de publicación de la revista.

Estructura.- El formulario tiene una extensión de 4 páginas y está constituido por 23 afirmaciones breves agrupadas en 4 categorías y distribuidas como sigue: prioridad y novedad del tema (2); proceso de investigación (14); estructura y estilo (4); recomendaciones sobre aceptación/rechazo (3).

La opinión que se recaba sobre los diversos items se vierte en un formulario de respuestas de elección múltiple los cuales varían entre 2 y 6 opciones, la última de las cuales adopta un formato abierto.

Observaciones.- Este formulario, -el cual se inauguró en 1988 con una segunda versión revisada en 1991- ha probado ser muy útil para el objetivo definido. Tiene las ventajas de ser breve, claro, directo y completo para los grandes rubros de la investigación. Presenta el inconveniente de no ahondar en aspectos del diseño, del análisis y de la interpretación de los datos. Además algunos árbitros se sienten «encasillados» por las respuestas estructuradas.

PAUTA N° 2

*Elementos para el análisis crítico de una investigación.*⁵²

Origen y objetivos.- Se trata del apéndice de un libro de texto sobre investigación en enfermería, de amplia difusión en los Estados Unidos. Su objetivo es formular preguntas relativas a la fase conceptual, a la fase empírica y a la fase interpretativa de la investigación. Las autoras del libro suponen que el lector de esta lista ha leído la obra y que está familiarizado con los conceptos y con la terminología. Por ello el coger la lista fuera de contexto puede disminuir su comprensión.

Estructura.- El formulario consta de 41 frases breves en forma de preguntas las cuales, en general, son unívocas, claras y directas; se distribuyen como sigue: el problema y los términos principales (4); la revisión de la literatura y el marco teórico (2); los objetivos e hipótesis (6); la población y la muestra (7); los procedimientos de recolección de datos (7); la presentación de datos (4); el análisis de datos (4); la discusión y conclusión (7).

Observaciones.- Esta lista de revisión tiene la ventaja de profundizar en el proceso de investigación con diferente grado de detalle. Aceptando el hecho de que el esquema puede quedar largo para evaluar algunas publicaciones y corto de preguntas para otras, en promedio ha probado ser una pauta útil. La forma de redacción de las preguntas induce a respuestas de «sí» o «no», lo cual no permite valores intermedios para aquellas variables de tipo continuo. Por otra parte diez de las preguntas incluyen dos o tres asuntos distintos; de hecho si se desglosaran estas preguntas el total subiría de 41 a 55.

La lista de preguntas calza con las etapas de la investigación e indirectamente con la estructura del cuerpo del artículo original. De hecho no hay preguntas sobre las partes preliminares (título, autores, resumen) ni de las partes finales (referencias, apéndices). Tampoco incluye preguntas sobre el estilo.

PAUTA N° 3

*Evaluación de comunicados de investigación.*⁵³

Origen y objetivos.- La propuesta que nos ocupa proviene del Capítulo 26 del libro que se señala en la referencia. La obra es un extenso libro de texto sobre investigación científica en ciencias de la salud de amplia utilización en los EUA, y en países de América Latina en la versión en idioma español que comentamos. Los 25 capítulos del libro que la preceden han examinado las bases conceptuales y metodológicas para mejorar la práctica de la investigación. En este capítulo se proporcionan pautas específicas para la revisión crítica e inteligente de la investigación.

Estructura.- Este formulario es el más extenso de todos cuantos conocemos, con 8 páginas las cuales albergan 169 items agrupados en 5 categorías siguiendo la estructura del artículo original: introducción (26); métodos (78); resultados (20); discusión (21); otros aspectos (24).

Observaciones.- Las numerosas preguntas permiten explorar algunos aspectos en mucho detalle. El hecho de que algunas preguntas sean sub-preguntas de las anteriores permite cierta profundidad. Además la forma de redacción induce a respuestas «sí» o «no»; valen aquí los comentarios de la pauta anterior. El 45% de las preguntas se refieren a la sección de material y métodos, lo cual se explica por el objetivo metodológico del libro del cual forma parte. La última agrupación «otros aspectos» se dedica al título, resumen, bibliografía, apéndices, estructura y estilo lo cual constituye un buen complemento del enfoque básicamente de diseño y de metodologías del extenso formulario.

PAUTA N° 4

*Evaluación crítica de la investigación publicada.*¹⁷

Origen y objetivos.- Esta pauta corresponde al último capítulo de un libro de texto sobre introducción a la investigación en las ciencias de la salud cuyos autores son profesores del departamento de ciencias del comportamiento en una universidad de Australia. El objetivo es demostrar cómo se pueden aplicar conceptos seleccionados sobre diseño, medida y datos estadísticos a la evaluación crítica de la investigación publicada.

Estructura.- La pauta se organiza en torno a la evaluación de los apartados específicos del cuerpo del artículo original. Su extensión alcanza 25 ítems agrupados en 4 categorías: introducción (4); métodos (12); resultados (3); discusión (6).

Observaciones.- Aún cuando el desglose de ítems aparece un tanto reducido, esta pauta incluye un rasgo único de gran valor práctico. Se trata de una tabla en dos columnas en que la de la izquierda enumera 24 problemas que se pueden identificar en el artículo y la de la derecha indica las posibles implicaciones de esos problemas.

PAUTA N°5

*Lista de evaluación en la investigación.*⁵⁵

Origen y objetivos.- Esta propuesta tiene su origen en el capítulo 10 del libro «Manual de Métodos de Investigación» de un autor británico. Se trata de una extensa y detallada lista para revisión de investigaciones publicadas, la cual cubre el proceso de comienzo a fin.

Estructura.- El autor propone una lista que incluye 77 ítems agrupados en 15 categorías con la siguiente distribución de frecuencias: comienzo (5); problema (4); revisión de la literatura (6); hipótesis o preguntas de investigación (8); población y muestra (6); diseño (8); métodos de recolección de datos (6); instrumentación (11); pruebas estadísticas y análisis (5); cronograma (2); personal (1); referencias (5); apéndices (2); general (4); varios (4).

Observaciones.- La primera categoría «comienzo» incluye preguntas referentes al título y al resumen; las tres categorías siguientes corresponden a la introducción del artículo; las siguientes cuatro forman parte del material y métodos; la categoría siguiente pertenece a resultados. Tal vez no se justifiquen categorías separadas para cronograma, personal, general y varios. Llama la atención que una lista tan extensa

y detallada en aspectos de metodologías resulte incompleta por carecer de preguntas sobre el importante capítulo de discusión y conclusiones.

PAUTA N° 6

La literatura médica ⁵⁴

Origen y objetivos.- Este es el capítulo 16 de un excelente texto de estadística para investigación médica. El autor comparte la conveniencia de disponer de una lista de puntos específicos al momento de leer un documento de investigación. Sin embargo argumenta que es imposible producir un conjunto de preguntas que sean adecuadas para analizar cualquier documento de investigación. En consecuencia el objetivo es proponer dos pautas diferentes: una para estudios generales y otra para ensayos clínicos.

Estructura.-

a) *Pauta general* la cual se compone de 17 preguntas distribuidas como sigue: diseño del estudio (8); conducción del estudio (1); análisis y presentación (6); apreciación global (2).

b) *Pauta para ensayo clínico* que consiste en una lista de 31 preguntas distribuidas en las mismas categorías que la pauta general: diseño del estudio (17); conducción del estudio (4); análisis y presentación (8); apreciación global (2).

El cuestionario se completa con tres columnas de medición cualitativa frente a cada pregunta: «Si», «Dudoso» y «No».

Observaciones.- El dividir la evaluación del estudio en una apreciación general y luego una especial para los ensayos clínicos utilizando dos pautas distintas constituye una propuesta muy razonable y atractiva. Además el incluir las 2 opciones extremas (si y no) más una opción intermedia (dudoso o no concluyente) es un elemento facilitador del proceso de escrutinio del texto del artículo original. La extensión de ambos cuestionarios es adecuada. El énfasis en el diseño, análisis y presentación apunta a aspectos que originan las faltas más frecuentes.

APÉNDICE I

EJERCICIO SOBRE REDACCIÓN DE TÍTULOS

M.A. Bobenrieth Astete

Título original: «Resultados de un estudio de quimioprofilaxis intravesical, prospectivo, doble aleatorio, entre dos drogas: La adriamicina y el mitomicin; y dos modos de iniciar las instilaciones: precoz y tardío. Efecto sobre la recidiva y la progresión».

Faltas: Demasiado extenso por falta de concisión; contiene numerosas palabras superfluas e inútiles (35 palabras); sobreexplicación («Resultados de un estudio de...»); mención innecesaria al diseño; uso de subtítulo; título «cortado» tres veces.

Título propuesto: «Quimioprofilaxis intravesical con adriamicina frente al mitocin C en el carcinoma vesical superficial» o bien:

«Quimioprofilaxis intravesical con adriamicina comparado con mitocin C en el carcinoma vesical superficial».

Título original: «¿Modifica el pronóstico del carcinoma renal la vía de abordaje utilizada para nefrectomía radical?»

Faltas: Título en tono interrogativo; imprecisión por uso de «pronóstico» en vez de «supervivencia».

Título propuesto: «Supervivencia del carcinoma renal según la vía de abordaje en la nefrectomía radical» o bien:

«Lunbotomía, laparotomía transperitoneal y toracofrenolaparatomía en la nefrectomía radical y supervivencia del carcinoma renal».

Título original: «¿Produce la RTU diseminación metastasica en el cáncer de próstata tratado hormonalmente?»

Faltas: Título en tono interrogativo; uso de abreviaturas.

Título propuesto: «Resección transuretral como factor de diseminación metastasica del cáncer de próstata tratado hormonalmente»

Título original: «Mitoxantrone (MXT) «versus» Mitomycin C (MMC) en el tratamiento ablativo de los tumores vesicales superficiales Ta, T1. Estudio prospectivo y al azar, en fase III».

Faltas: Uso de abreviaturas; demasiado extenso por falta de concisión; uso de barbarismos (versus); subtítulo; mención del diseño.

NOTA: Versus. El uso de la preposición latina «versus», «contra» entre dos nombres de personas, colectividades o cosas (ej. ciudad versus campo; academia versus servicios, etc.) es un anglicismo, o más bien un americanismo. Se la considera un esnobismo, es decir, la imitación con afectación de maneras, opiniones, etc. de personas, grupos o culturas a quienes se considera distinguidos. En español correcto se dice «frente a» o «comparado con».

Título propuesto: «Mitoxantrone comparado con Mitomycin C en el tratamiento ablativo de tumores vesicales superficiales» o bien:

«Tratamiento ablativo de tumores vesicales superficiales con Mitoxantrome comparado con Mitomycin C».

Título original: «Profilaxis antimicrobiana en cirugía endoscópica urológica: Pefloxacino versus Ceftriaxona».

Faltas: Título cortado; barbarismo («versus»)

Título propuesto: «Pefloxacino comparado con Ceftriaxona en la prevención de infecciones en cirugía endoscópica urológica».

Título original: «Actividad sexual tras la cirugía de la hipertrofia prostática benigna».

Faltas: Sintaxis defectuosa.

Título propuesto: «Cirugía en la hipertrofia prostática benigna y actividad sexual» o bien: «Efecto de la cirugía sobre la actividad sexual en la hipertrofia prostática benigna».

Título original: «Tratamiento de los tumores de próstata con termocoagulación intersticial con láser de neodimio-YAG (Un tratamiento nuevo dentro de la cirugía mínimamente invasiva)».

Faltas: Subtítulo entre paréntesis, de contenido superfluo e innecesario.

Título propuesto: «Tratamiento de tumores de próstata con termocoagulación intersticial con laser de neodimio-YAG».

Título original: «Papel de la testosterona en las variaciones microcirculatorias testiculares producidas en la rata por la administración de altas dosis de gonadotropina coriónica humana».

Faltas: Imprecisión y falta de concisión.

Título propuesto: «Papel mediador de la testosterona en las variaciones microcirculatorias testiculares producidas por la gonadotropina coriónica humana en ratas»

Título original: «Eficacia diagnóstica del TC y RM en el estadiaje local de los tumores renales».

Faltas: Uso de abreviaturas.

Título propuesto: «Tomografía computarizada comparada con resonancia magnética en el estadiaje local de tumores renales».

Título original: «Controversia ante la actitud terapéutica en el adenocarcinoma de estadio B2, C y D1».

Faltas: Encabezamiento del título con el término «controversia» sin dar luces sobre los aspectos claves de la controversia que se presentan en el estudio.

Título propuesto: «Tratamiento del adenocarcinoma de estadio B2, C y D1».

Título original: «Estudio prospectivo y randomizado de la respuesta completa, índice de recidivas y progresión en el carcinoma vesical superficial tratado con Mitomycina-C sola versus Mitomycina-C y BCG alternos».

Faltas: Demasiado extenso; sobreexplicación (estudio); mención metodológica innecesaria; barbarismo por uso de la palabra «versus» en lugar de «comparado con».

Título propuesto: «Tratamiento del carcinoma vesical superficial con Mitomycina-C comparado con Mitomycina-C y BCG alternos».

Título original: «Carcinoma renal. Nuestra experiencia de 20 años».

Faltas: Estilo telegráfico; subtítulo irrelevante con personalización de la idea («nuestra»).

Título propuesto: «Diagnóstico, tratamiento y supervivencia del carcinoma renal en Santander, 1970-1989».

Título original: «Flurbiprofen: Alternativa terapéutica en el dolor del cólico nefrítico»

Faltas: Demasiado corto e incompleto; título cortado.

Título propuesto: «Flurbiprofen comparado con dipirona más hioscina en el dolor del cólico nefrítico».

Título original: «Estudio clínico, epidemiológico y microbiológico de infección urinaria en pacientes con trasplante renal en un centro especializado de Argentina».

Faltas: Sobreexplicación («estudio»); uso demasiado amplio del término «epidermiológico».

Título propuesto: «Infección urinaria en pacientes con trasplante renal en un centro especializado de Argentina».

Título original: «Estudio clínico-patológico del carcinoma de células renales».

Faltas: Sobreexplicación («estudio»); incompleto.

Título propuesto: «Anatomía patológica y supervivencia en carcinoma de células renales».

Título original: «Tratamiento de los tumores vesicales superficiales recidivantes con laser Nd:Yag: un año de experiencia».

Faltas: Título partido; subtítulo; referencia innecesaria a la duración del estudio (un año de experiencia).

Título propuesto: «Laserterapia en tumores vesicales superficiales recidivantes en un hospital de Madrid».

Título original: «Síndrome de prostatismo en consultas externas de urología. Su incidencia en un hospital comarcal». **Faltas:** Sobreexplicación (síndrome de); subtítulo; uso discutible del término «incidencia».

Título propuesto: «Prostatismo en la consulta externa de urología en el Hospital Comarcal de Denia».

Título original: «Estudio doble ciego de la eficacia analgésica en el cólico nefrítico de la asociación de dipirona y espasmolítico con el ketorolaco trometamol».

Faltas: Sobreexplicación («estudio»); indicación innecesaria del diseño («doble ciego»). **Título propuesto:** «Analgésia con dipirona más espasmolítico comparado con ketorolaco trometamol en el cólico nefrítico».

Título original: «Valoración del Prazosin en el tratamiento de la hipertrofia benigna de próstata. Estudio caso-control».

Faltas: Sobreexplicación («valoración de»); subtítulo; mención innecesaria del diseño (caso-control).

Título propuesto: «Tratamiento de la hipertrofia benigna de próstata con Prazosin».

APÉNDICE II

EL PROCESO DE LA REVISIÓN BIBLIOGRÁFICA

Resumen esquemático

*Irene Fernández Romacho **

A. Introducción

Para que una investigación sea útil, sus hallazgos han de ser una prolongación de conocimientos y teorías previas, por tanto, el investigador debe estar actualizado en cuanto a lo que ya se sabe acerca de su tema de estudio y sobre cómo se ha logrado este conocimiento. Existen dos maneras de llevar a cabo esta actualización:

1. Estudio de antecedentes concretos, casos prácticos determinados, que requieren un sondeo de campo, para lo que serán necesarias la interrogación y la observación.
2. Revisión bibliográfica. Procedimiento que se lleva a cabo para:
 - a. Investigar sobre un tema
 - b. Estudiar y resolver problemas que se originan en casos prácticos
 - c. Para transmitir nueva información a alumnos

B. Objetivos de la revisión bibliográfica

1. Responder a cuestiones relacionadas con:
 - a. Información teórica
 - b. Información práctica
 - c. Información sobre diseños y sus correspondientes metodologías
2. Proporcionar al interesado las claves del estado de la cuestión
3. Dar al investigador las claves para no repetir lo que ya se sabe

C. Momentos en que ha de llevarse a cabo la revisión

El proceso de revisión bibliográfica, más o menos específico dependiendo del caso concreto, debe realizarse en diferentes momentos del proceso de investigación:

* Licenciada en Biblioteconomía y Documentación. Universidad de Granada

1. Antes de concretar el tema, para ayudarse a encauzar la investigación. La revisión será opcional en esta instancia, según el nivel de claridad y concisión que tenga el investigador acerca de su tema de estudio.
2. Una vez identificado el tema de estudio, para posicionarse en el estado actual de la cuestión.
3. En la interpretación y comparación de resultados, donde se revisará de manera más selectiva lo publicado para confrontarlo con los resultados obtenidos con otros métodos de investigación: recogida de datos, pruebas de campo, etc.

D. Búsqueda y recuperación de la información

Cuando el investigador ha concretado lo que va a investigar, ha de conseguir información al respecto. La búsqueda y recuperación de la información consiste en extraer de una o varias colecciones de documentos aquéllos que se ajusten con más precisión al tema de interés. Tal como muestra el diagrama anexo, el investigador debe comenzar la búsqueda bibliográfica y la recuperación siguiendo estas fases:

1. Identificación de términos: el investigador debe saber de qué forma han denominado otros las mismas cuestiones que él intenta abordar. Para ello debe determinar los conceptos e identificar sus sinónimos. Una forma de ayudarse en esta tarea, aparte de indagar en sus propios conocimientos, es la consulta a obras de referencia relacionadas con el tema.

2. Búsqueda bibliográfica y recuperación de información: Una búsqueda perfecta sería aquella que daría lugar a la recuperación de todos los documentos (ni uno más ni uno menos) que el investigador elegiría, si pudiera examinarlos uno a uno. Frente a esta imposibilidad, el investigador trabaja en aproximaciones sucesivas ayudado de los diferentes sistemas de información existentes al respecto. Sobre estos sistemas las tareas a realizar son las siguientes:

- a. Búsqueda en fuentes secundarias. Fuente secundaria es aquel documento impreso o informatizado que contiene información referencial (información mediata o indirecta), datos reelaborados sobre fuentes de primera mano. El investigador interrogará a los diferentes sistemas de información para que éstos le proporcionen la referencia de documentos originales que pueden interesarle.
- b. Identificación de fuentes primarias y extracción de referencias. Fuentes primarias son documentos originales, de primera mano, que contienen datos inmediatos y directos. El investigador deberá reconocer los que le interesen partiendo de los datos identificativos que el sistema de información le ha proporcionado al efectuar su consulta. Debe extraer esas referencias bibliográficas, tomar nota de sus datos para volver a abordar al sistema en la fase siguiente.
- c. Localización de las referencias en los catálogos de los sistemas de información: Deberá buscar y reconocer los documentos seleccionados durante la búsqueda en los catálogos de los diferentes sistemas de información, donde es probable que las referencias sean más extensas y precisas.

- d. Selección y expurgo de las fuentes primarias: Una vez localizados en el catálogo los documentos escogidos, deben seleccionarse aquellas publicaciones que se estudiarán más detenidamente y serán la base de la investigación. Para ello, deben evaluarse los documentos teniendo en cuenta una serie de características que pueden extraerse de los datos que aparezcan en su referencia bibliográfica, si ésta es completa. Las cuestiones a tener en cuenta son:

- Los textos, la autenticidad y la originalidad.
- La prioridad de los documentos de primera mano, frente a sus traducciones.
- La preferencia de los originales frente a las copias.
- Preferencia por las últimas ediciones, es decir, aquellos documentos que hayan sido revisados y renovados por sus autores.
- Hay que tener en cuenta si son o no ediciones críticas, anotadas.

E. Lectura crítica y evaluación de las fuentes

La lectura de las fuentes primarias y su evaluación constituyen la última fase de la revisión bibliográfica e influyen directamente sobre los resultados de la investigación.

1. Lectura

- a. Criterios de lectura. En la lectura intervienen determinados criterios que pueden tomarse como correlativos a la hora de realizar esta tarea:
- Percepción: reconocer las letras o signos procurando que el campo visual sea lo más amplio posible.
 - Comprensión: entender el sentido de las ideas que da el autor.
 - Asimilación: clarificar las ideas dentro del conocimiento intelectual y científico, ponderarlas o discernirlas e incorporarlas a los conocimientos previos.
 - Evaluación: interpretar y criticar lo leído, es decir, desvelar ideas implícitas y explícitas y enjuiciar el escrito desde el punto de vista gramatical, la lógica (referida al orden en el pensamiento) y el contenido material o doctrinal.
- b. Tipos de lectura
- Lectura de corrido, cuyo fin es obtener información de tipo general y se realiza atendiendo al título, los datos del autor (trayectoria profesional, etc.), el resumen, el índice, los objetivos, la conclusión, los hallazgos, la introducción y la bibliografía.
 - Lectura detenida, que incluye atención, reflexión y toma de notas.
 - Lectura de control, que consiste en una lectura detenida centrada en descubrir ideas nuevas.

c. Técnicas de lectura

- Realizar una lectura rápida para tomar una idea general y seleccionar los apartados que debemos leer con más detalle.
- Realizar una lectura detenida de los apartados seleccionados, intentando comprender antes de aceptar o refutar, y anotar ideas claves al margen.
- Llevar a cabo una reflexión global y crítica sobre el contenido.

En general, leeremos más deprisa si:

- Nos centramos en los aspectos más importantes del texto, que se encuentran habitualmente al principio y final de los párrafos, resaltados tipográficamente, y después de expresiones conectivas como “por lo tanto”, “en efecto”, “en definitiva”, etc.
- Leemos párrafos más que frases.
- Leemos frases más que palabras.
- Somos rápidos en pasar de un renglón a otro.
- Entendemos el significado y no titubeamos en el reconocimiento de signos.
- No volvemos atrás.
- No vocalizamos.
- No pronunciamos mentalmente lo leído.

2. Evaluación

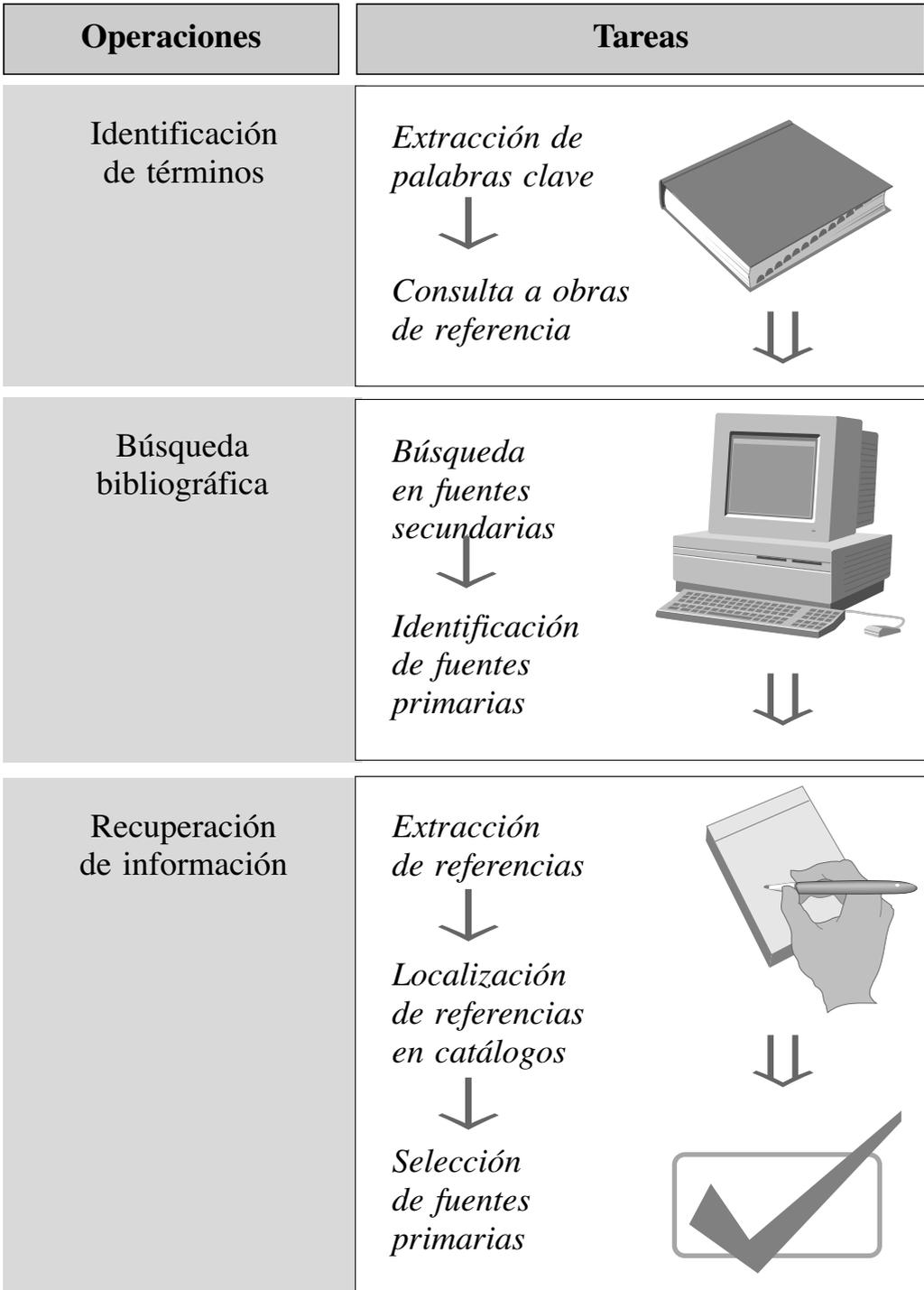
La evaluación de las fuentes intenta determinar el valor intrínseco de la información (validez, calidad y viabilidad) relacionado con los objetivos, el diseño, la metodología, la presentación de los resultados, el análisis, la interpretación y la conclusión.

a. Criterios de evaluación

- Criterio de integridad: analiza la cabalidad de la publicación y cuán completa es.
- Criterio de consistencia: analiza la solidez y coherencia del texto.
- Criterio de valor práctico: determina la utilidad de la fuente analizada.

b. Principios de evaluación

- Pertinencia: ver si la información viene a propósito del tema de estudio.
- Seguridad en el contenido: valorar la capacidad, método y experiencia del autor, la editorial, el traductor, el estilo en la redacción, la claridad, etc.
- Vigencia (actualidad en la información): constatar la fecha de registro de la propiedad intelectual y de las referencias bibliográficas que contiene.
- Objetividad: comprobar si se exponen con imparcialidad los distintos puntos de vista que se publican sobre un problema.
- Normalización: constatar que la disposición de los epígrafes y el cuerpo de la obra coincide con lo anunciado en la introducción y en el índice.



Esquema IV.2. Búsqueda y recuperación de la información. Fuente: Elaboración propia

BIBLIOGRAFÍA

Bobenrieth Astete, MA. *El artículo científico original. Estructura, estilo y lectura crítica*. Granada: Escuela Andaluza de Salud Pública, 1994.

Guía práctica para la elaboración de un trabajo científico. Bilbao: CITA, 1994. *Fundamentos de Información y Documentación*. Madrid: EUDEMA, 1989.

Pinto Molina, M. "El análisis documental". En: *Fundamentos de Información y Documentación*. Madrid: EUDEMA, 1989 (págs. 263-279).

-"El análisis de contenido: la operación de resumir". En: *Fundamentos de Información y Documentación*. Op. cit. (Págs. 385-413).

Polit DF, Hungler BP. *Investigación científica en Ciencias de la Salud*. 4ª Ed. México: Interamericana McGraw-Hill, 1994.

APENDICE III
PAUTAS PARA LA EVALUACION
DE ARTICULOS ORIGINALES

M.A. Bobenrieth Astete

FORMULARIO DEL ARBITRO *

*Boletín de la Oficina Sanitaria Panamericana/
Bulletin of the Pan American Health Organization*

Manuscrito:.....

Autor(es):

Arbitro:

Fecha:

1. La prioridad del tema para la revista es

- alta
- mediana
- baja
- nula

2. El problema general

- se identifica de forma inmediata y clara
- no se describe en cuanto a origen, magnitud e importancia
- no se identifica
-

3. El problema específico de la investigación

- se delimita con claridad
- no se define con precisión
- no se distingue del problema general de la investigación
-

4. La revisión bibliográfica

- expone los trabajos publicados más importantes y recientes relacionados con el tema
- no es actual o suficiente
-

* HBI/ES 1991, 2ª versión

5. Los objetivos y las hipótesis

- son claros, precisos, factibles y medibles
- son demasiado generales, vagos o ambiguos
-

6. El diseño

- es apropiado para el objetivo del estudio
- no es adecuado
- no está descrito suficientemente
-

7. Los métodos y las técnicas

- son apropiados
- son inapropiados
-

8. Los datos sobre materiales y métodos

- son suficientes para repetir el estudio
- son insuficientes
-

9. Las pruebas estadísticas

- son apropiadas para los tipos de variables utilizadas y responden a las hipótesis formuladas
- son inadecuadas
- no están descritas con suficiente detalle
- no están descritas
-

10. Los resultados

- presentan información pertinente a los objetivos del estudio
- no son pertinentes
- incluyen detalles suficientes para justificar las conclusiones
- son insuficientes
-

11. Las tablas

- son suficientes y apropiadas
- deben mejorarse para presentar e interpretar los datos con mayor claridad
- son excesivos y redundantes
- la(las) tablas(s) número(s) puede(n) suprimirse
- la tabla número y la pueden combinarse
-

12. Las ilustraciones (gráficos, fotografías, etc.)

- son suficientes y apropiadas
- deben mejorarse para presentar e interpretar los datos con mayor claridad
- son excesivas y redundantes
- la(s) figura(s) número puede(n) suprimirse
- es aconsejable agregar una figura para ilustrar
- los originales son de calidad deficiente
-

13. La discusión

- examina e interpreta los resultados correctamente en función de su significado y de sus limitaciones
- no interpreta los resultados
- adolece de omisiones importantes
- contiene generalizaciones excesivas
- repite los resultados
-

14. La conclusión

- se justifica porque se apoya en los hallazgos
- no se justifica porque no se apoya en los hallazgos
-

15. Las referencias bibliográficas

- son suficientes y apropiadas
- son insuficientes
- no son actuales
- no se mencionan referencias imprescindibles; por ejemplo:

16. La información que contiene el artículo

- es nueva y representa un aporte valioso
- repite resultados ya conocidos
- incluye material irrelevante

17. El título

- es claro, conciso e informativo
- debe modificarse. Se propone:

18. El resumen

- es claro y completo (incluye los objetivos, el diseño, las circunstancias en que se realizó, la descripción de los participantes, los métodos, las variables medidas, los resultados principales y las conclusiones más importantes)
- es ambiguo e incompleto. Debe reescribirse

19. La estructura

- es apropiada
- es impropia y debe cambiarse
-

20. El estilo

- es apropiado; tiene claridad, concisión y secuencia lógica
- es escueto, poco claro o deficiente
- es redundante y debe condensarse
-

21. El artículo

- es excelente
- es bueno
- es regular
- es malo
- es pésimo

RECOMENDACION**22. El artículo:**

- puede publicarse tal como está
- podría publicarse con modificaciones mínimas
- podría publicarse con modificaciones sustanciales
- debe rechazarse; no es idóneo para el Boletín de la OPS/Bulletin of PAHO
- El autor(s) podría someterlo para publicación a otra revista. Se sugiere:

23. OTROS COMENTARIOS: (agradeceremos mecanografiar si es posible)

Firma:

ELEMENTOS PARA EL ANALISIS CRITICO DE UNA INVESTIGACION *

A. Fase conceptual

Problema y términos

1. ¿Se plantea el problema general con prontitud y claridad?
2. ¿Se delimita el problema general en un problema específico de investigación?
3. ¿Se fundamenta el problema con una base de antecedentes adecuados y necesarios para el estudio?
4. ¿Se definen los términos principales desde el punto de vista operacional?

Revisión de la literatura y marco teórico

5. ¿La revisión bibliográfica es pertinente, suficiente y actualizada?
6. ¿El marco teórico provee una base conceptual que permita relacionar teorías que inclinen hacia un enfoque de investigación?

Objetivos e hipótesis

7. ¿Corresponden los objetivos y/o hipótesis directamente al problema de investigación?
8. ¿Los objetivos y las hipótesis son claros, precisos, factibles y medibles?
9. ¿Es explícita la justificación teórica de la hipótesis?
10. ¿Las hipótesis son específicas para una relación entre variables de manera que la investigación pueda confirmarlas o rechazarlas?
11. ¿Se plantean limitaciones al estudio?
12. ¿Está clara la manera en que el estudio ampliará y mejorará resultados anteriores publicados?

B. Fase empírica

Población y muestra

13. ¿El diseño del estudio es apropiado para el objetivo del mismo?
14. ¿Son suficientemente específicos los parámetros de la población objeto de estudio
¿Queda claro a qué población podría generalizarse los resultados?
15. ¿Es la muestra representativa del universo (o población) definida?

* Traducción y adaptación de Manuel A. Bobenrieth de la obra: Downs, Florence; Newman Margared «Elements of a Research Critique». In *A Source Book of Nursing Research*. Third Edition. Philadelphia: F.A. Davis Co, 1983

16. ¿Es apropiado el método de selección de la muestra?
17. ¿Es apropiado el tamaño de la muestra? ¿Como se ha substanciado?
18. ¿Introduce el método de selección de la muestra algún error o sesgo?
19. ¿Qué pruebas indican que se han protegido los derechos de los sujetos?

Procedimiento de acopio de datos. Instrumentos

20. ¿Se plantea el grado de confiabilidad de las fuentes de información, tanto documentales como de campo?
21. ¿Son apropiados para el estudio los métodos de acopio de datos? ¿Permiten efectivamente obtener los datos que busca el investigador?
22. ¿Especifica el investigador la validez y la fiabilidad de los instrumentos empleados?
23. ¿Qué medidas se han tomado para controlar las variables externas?
24. ¿Son reproducibles las condiciones de la intervención?
25. ¿Es la variable independiente lo bastante fuerte para resistir influencias concomitantes?
26. ¿Es la variable independiente empleada representación válida de la realidad, en un estudio de diseño experimental?

Presentación de datos

27. ¿Los resultados presentan la información pertinente a los objetivos del mismo?
28. ¿El texto es claro, conciso y secuencial?
29. ¿Las tablas son necesarias, suficientes y apropiadas?
30. ¿Las ilustraciones (figuras, etc..) son necesarias, suficientes y apropiadas?

C. Fase interpretativa

Análisis de datos

31. ¿Los métodos y las técnicas son apropiados?
32. ¿Se exponen claramente los análisis en relación a cada objetivo o hipótesis?
33. ¿Está claro qué prueba estadística se utilizó y qué valores se obtuvieron?
34. ¿Las pruebas estadísticas son apropiadas para los tipos de variables utilizadas y para responder a las hipótesis formuladas?

Discusión y conclusiones

35. ¿La discusión examina e interpreta los resultados correctamente en función de su significado y de sus limitaciones?
36. ¿Ha relacionado el investigador los resultados con el marco teórico del estudio?
37. ¿Logra el investigador convencer al lector sobre la validez interna y la validez externa del estudio?
38. ¿Ha generalizado demasiado el investigador?
39. ¿La conclusión responde a los objetivos, apoyándose en los hallazgos?
40. ¿Cuáles son las consecuencias prácticas de los resultados?
41. ¿Se sugieren otras investigaciones complementarias?

EVALUACIÓN DE COMUNICADOS DE INVESTIGACIÓN *

EVALUACIÓN DE LA INTRODUCCIÓN

Problema

1. ¿El problema se plantea en forma clara y precisa?
2. ¿El problema es demasiado grande o complejo como para resolverse en una sola investigación?
3. ¿El planteamiento aporta información precisa respecto a las variables independiente y dependiente?
4. ¿La pregunta o duda puede resolverse con pruebas empíricas?
5. ¿Se incluyen definiciones de términos necesarios para el entendimiento claro del estudio?
6. ¿Se comenta la importancia del problema para la asistencia?
7. ¿Hay posibilidad de que el problema tenga importancia más allá de un terreno local?

Revisión de literatura

8. ¿El comunicado vincula el problema con investigaciones previas sobre el tema?
9. ¿Existe dependencia excesiva de fuentes secundarias cuando se han obtenido las fuentes primarias?
10. ¿Se omiten referencias importantes?
11. ¿La revisión incluye los comunicados recientes en la literatura?
12. ¿Se pone énfasis excesivo a artículos de opinión y se concede poca importancia se concede a trabajos empíricos?
13. ¿La redacción es adecuada?. ¿Incluye demasiadas citas de las fuentes originales?
14. ¿La revisión es simplemente una condensación de trabajos pasados?. ¿Evalúa críticamente y compara las contribuciones de los estudios fundamentales?
15. ¿La revisión se organiza de forma tal que el desarrollo de las ideas es claro?
16. ¿La revisión concluye con una sinopsis breve de la literatura y sus consecuencias respecto al problema por investigar?

Marco conceptual (marco teórico)

17. ¿El comunicado intenta vincular el problema con un marco o estructura teórico o conceptual?
18. ¿El marco teórico se vincula con el problema en forma natural o por el contrario no hay una vinculación adecuada?

* Polit DF, Hungler B.P. *Investigación Científica en Ciencias de la Salud*. Tercera Edición. Capítulo 26. Mexico: Interamericana/Mc Graw-Hill, 1991

19. ¿Sería más adecuado otro marco conceptual?
20. ¿Son lógicas las deducciones de una teoría o marco conceptual?

Hipótesis

21. ¿El comunicado identifica las hipótesis que se someten a prueba?
22. ¿Cada hipótesis expresa una relación predicha entre dos o más variables?
23. ¿Las hipótesis siguen un curso lógico a partir de la base teórica o revisión de la literatura?. De no ser así, ¿qué justificación se da a las predicciones del investigador?
24. ¿Todas las hipótesis se pueden corroborar?
25. ¿Las hipótesis se plantean en forma concisa y clara?
26. ¿Las hipótesis indican la población general de interés?

EVALUACION DE LOS METODOS

Sujetos

27. ¿Se identifica y describe la población?
28. ¿El comunicado indica si se estudió la población completa, o sólo una muestra?
29. ¿Se describen claramente las técnicas de selección de la muestra?
30. ¿El diseño del muestreo puede producir una muestra representativa?
31. ¿Si el diseño del muestreo es relativamente débil, como el caso del muestreo no probabilístico, ¿se identifican los posibles errores sistemáticos o parcialidades de las muestras?
32. ¿Se describen las características de tamaño y otras fundamentales de la muestra?
33. ¿Es el tamaño de la muestra lo suficientemente grande ante la heterogeneidad de la población?
34. ¿El comunicado indica el índice de respuestas, esto es, el porcentaje de personas que fueron entrevistadas y que aceptaron participar en el estudio?

Métodos de obtención de datos: aspectos generales

35. ¿Se identifican y describen con nitidez los instrumentos para la obtención de los datos?
36. ¿Los métodos de obtención de datos constituyen la forma más adecuada para medir las variables críticas?
37. Si se crea específicamente un instrumento para el estudio, ¿se describen las técnicas para su creación?
38. Si los instrumentos son nuevos o se han adaptado de otros anteriores, ¿se hace una prueba preliminar adecuadamente?
39. ¿Se presentan pruebas de la confiabilidad de los instrumentos?
40. De haber una respuesta afirmativa, ¿son los coeficientes de confiabilidad de magnitud aceptable?; ¿el tipo de estimados de confiabilidad obtenido es el más adecuado?
41. ¿Se presenta prueba alguna de la validez de los instrumentos?
42. En caso afirmativo, ¿las pruebas indican que los instrumentos son lo suficientemente válidos para el uso al que se les destinó?
43. ¿El tipo de validez expuesto, (esto es, contenido, criterio, constructo) es el más relevante para los instrumentos en consideración?

Métodos de obtención de datos: Técnicas específicas

Técnicas observacionales

44. ¿Se definen con claridad las conductas o circunstancias por observar?
45. ¿Los fenómenos por observar son los mismos que se describen en el planteamiento del problema?
46. ¿La unidad de conducta adoptada es apropiada para el problema en estudio?
47. ¿El grado de estructura usado en la observación es congruente con los fines del estudio?
48. ¿El grado de ocultamiento utilizado es congruente con los fines del estudio y los principios éticos?
49. ¿El sistema de categorías, si lo hay, se describe adecuadamente y es completo?
50. ¿Se necesita que los observadores hagan una cantidad excesiva de inferencias?
51. ¿Se pide a los observadores que codifiquen demasiadas conductas complejas en muy poco tiempo?
52. ¿Se expone el método de muestrear conductas, y éste es apropiado?
53. ¿Los observadores se han preparado lo suficiente para usar los métodos de observación?
54. ¿Se expone la confiabilidad entre uno y otro observador, y es lo suficientemente grande?
55. ¿Se exponen los posibles vicios o parcialidades, que provienen del observador o del observado?
56. ¿Se expone el problema de la reactividad del sujeto y sus consecuencias?

Planes y cuestionarios de la entrevista

57. ¿En el comunicado se describen o incluyen los planes?
58. ¿Las preguntas cubren adecuadamente las complejidades del problema por estudiar?
59. ¿Se usan eficazmente y en las proporciones apropiadas las preguntas abiertas y cerradas?
60. ¿La redacción de las preguntas es sencilla y clara?
61. ¿Las preguntas tienden a viciar las respuestas en una dirección determinada?
62. ¿El plan tiene el tamaño apropiado?
63. ¿Son claras las instrucciones para los entrevistadores o los respondientes?
64. ¿Se describen la preparación y experiencia de los entrevistadores?
65. ¿El comunicado indica el sitio en que ocurrieron las entrevistas?
66. ¿Es significativo y apropiado el ordenamiento de las preguntas en el plan?
67. En preguntas cerradas, ¿las respuestas cubren adecuadamente las alternativas?
68. ¿Se aseguró a los participantes la confidencialidad o el anonimato?
70. En el caso de cuestionario, ¿se describen las técnicas de vigilancia y, de ser así, son adecuadas?

Escalas

71. ¿Se explica la razón para escoger una técnica de cuantificación escalar, a diferencia de otra (como serían las escalas de Likert en comparación con las de Guttman)?
72. ¿El método para colocar puntos reactivos en una escala es empírico o por criterios (esto es, se hace por análisis factorial o por análisis de reactivos)?
73. ¿Se describen los métodos para eliminar o llevar al mínimo los grupos de respuestas (respuestas estereotipadas)?
74. ¿Se equilibran los puntos negativos con los positivos?

75. ¿La escala es unidimensional?
76. ¿Son claras las instrucciones para los respondientes?
77. Si se usó una escala preexistente, ¿se explicó su importancia en razón de los objetivos del estudio?

Diseño de la investigación: aspectos generales

78. ¿Se describe adecuadamente el diseño del estudio?
79. El enfoque general —esto es experimental, cuasi—experimental o no experimental— es el mejor para someter a comprobación las hipótesis del estudio?
80. ¿El diseño controla los elementos que menoscaban la validez interna del estudio?
81. ¿Qué métodos se usan para controlar las diferencias individuales?
82. ¿Estos métodos son lo suficientemente eficaces?
83. ¿Se presta consideración adecuada, en el diseño, a las amenazas que pueden menoscabar la validez interna del estudio?
84. Si se utiliza un diseño transversal, ¿sería más apropiado un diseño longitudinal?
85. Si se usa el diseño longitudinal ¿se exponen los métodos para entrar de nuevo en contacto con los sujetos y resolver los problemas de deserción (atracción)?
86. ¿El diseño comprende todas las hipótesis por investigar?

Diseño de la investigación: aspectos específicos

Estudios experimentales

87. ¿Se describe el método para asignar, de manera aleatoria, los sujetos a grupos?
88. ¿Se describe el diseño en forma exacta y éste es apropiado?
89. ¿Se especifica la base para la selección del diseño?

Estudios cuasi-experimentales y preexperimentales

90. ¿Se describe el método para la selección de los grupos de comparación si los hay?
91. ¿Se siguen algunos casos para asegurar la equivalencia de los grupos de comparación, y son adecuados los pasos?
92. ¿El comunicado indica el grado de equivalencia de los grupos de comparación?
93. Si no se utiliza grupo de comparación, ¿se da alguna explicación de tal hecho?
94. Si se usa grupo de comparación, ¿se utilizan métodos para facilitar la interpretación de los resultados, y éstos son apropiados y satisfactorios?

Estudios no experimentales

95. ¿Se describe el método para la selección de los grupos de comparación si los hay?
96. ¿Se siguen algunos pasos, como el igualamiento o emparejamiento, para producir grupos de comparación equivalentes, y tales pasos son adecuados?
97. ¿El comunicado indica el grado con el cual los grupos de comparación utilizados (si los hay) son equivalentes?
98. ¿Se describen claramente los atributos o experiencias que diferencian a los grupos de comparación si los hay?
99. ¿Se describen las variables extrañas que no fueron controladas y que son importantes?

Técnicas de investigación

100. ¿Se describen con claridad y detalle las técnicas utilizadas para llevar a la práctica el diseño, de manera que permita su repetición por otro investigador?
101. ¿Se describen las técnicas para asegurar la constancia de las condiciones?
102. ¿Se exponen las formas de evitar la contaminación entre los grupos en tratamiento?
103. ¿Son convenientes las circunstancias de estudio (campo, en comparación con laboratorio) para el planteamiento por investigar?
104. ¿Se protegen los derechos de los sujetos?

EVALUACIÓN DEL ANÁLISIS Y LOS RESULTADOS

Técnicas para procesamientos de datos

105. ¿Los datos se analizaron por computadora o manualmente?
106. Si su análisis se hace por computadora, ¿el comunicado especifica el programa o paquete particulares utilizados?
107. ¿El comunicado describe las técnicas de codificación usadas?

Análisis de datos: aspectos generales

108. ¿Los datos se analizan por técnicas cualitativas, cuantitativas, o por ambas?
109. ¿El tipo de análisis es adecuado para los datos reunidos?

Estadística descriptiva

110. ¿Se exponen las técnicas de estadística descriptiva?
111. ¿La estadística descriptiva es adecuada para procesar los datos?
112. ¿La estadística descriptiva se presenta en forma coherente?
113. Si solamente se utiliza estadística descriptiva, ¿es suficiente para los fines de la investigación, o también hay que incluir la estadística inferencial?

Estadística inferencial

114. ¿Se utilizan pruebas de significación estadística para someter a prueba las hipótesis?
115. ¿Se utilizan pruebas paramétricas cuando se violan flagrantemente las suposiciones para usarlas?
116. ¿Se utilizan pruebas no paramétricas cuando tal vez sería conveniente una prueba paramétrica más potente?
117. ¿Se desperdicia información de manera innecesaria al transformar medidas de un nivel relativamente grande de medición, a otra de nivel menor (por ejemplo, transformar la talla en pulgadas, a la dicotomía estatura alta/baja)?
118. ¿Los estudios de significación son adecuados para corroborar las hipótesis de investigación?
119. ¿El comunicado indica el valor del estadístico calculado, el número de grados de libertad y el nivel de significación?

Datos estadísticos

120. ¿Se presentan con claridad los resultados?
121. ¿Se organiza con propiedad la presentación de los hallazgos?
122. ¿Las tablas y las figuras se usan eficazmente para esclarecer y exponer directamente los resultados?
123. ¿Se organizan las tablas y las figuras? ¿Se marcan apropiadamente y es fácil su comprensión?
124. ¿Hay algún signo de error (o de sesgo) en el comunicado de los hallazgos?

EVALUACION DE LA DISCUSION**Interpretación de los hallazgos**

125. ¿Se exponen todos los resultados importantes?
126. ¿Se interpreta cada resultado en término de la hipótesis original con la que se relaciona?
127. ¿Cada resultado se interpreta a la luz de hallazgos de estudios semejantes a la investigación?
128. ¿Se mencionan otras explicaciones de los datos y se expone la base de su rechazo?
129. ¿Las interpretaciones dan debida consideración a las limitaciones de los métodos de estudio?
130. ¿Las interpretaciones son congruentes con los resultados?
131. ¿Se hacen interpretaciones de causalidad no justificadas?
132. ¿Es lo suficientemente atractivo el lenguaje utilizado para comentar las interpretaciones?
133. ¿Las interpretaciones se organizan en una forma significativa?
134. ¿Hay algún signo de «vicios sistemáticos» o errores en las interpretaciones?
135. ¿La interpretación diferencia entre significación práctica y significación estadística?

Consecuencias

136. ¿Se pasan por alto las consecuencias del estudio, aunque se advierte que existe una base para ellas?
137. ¿Las consecuencias del estudio se exponen en términos de retención, modificación o rechazo de un esquema teórico/conceptual?
138. ¿Se describen las consecuencias de los hallazgos para la práctica profesional?
139. ¿Son apropiadas las implicaciones expuestas?
140. ¿Se hacen generalizaciones que no se justifican con base en la muestra utilizada?
141. ¿Se presta debida consideración a las limitaciones del estudio al comentar las consecuencias?

Recomendaciones

142. ¿Se hacen recomendaciones respecto a la forma de mejorar los métodos de estudio?
143. ¿Se hacen recomendaciones para tareas específicas, con base en las consecuencias?
144. ¿Se hacen redomendaciones para futuras investigaciones?
145. ¿Las recomendaciones son congruentes con los hallazgos y con los resultados afines a la investigación?

EVALUACION DE OTROS ASPECTOS DEL COMUNICADO

Título

146. ¿El comunicado tiene un título?
147. ¿El título tiene el tamaño razonable?
148. ¿En el título se mencionan las variables importantes?
149. ¿El título sugiere la población en investigación?

Resumen

150. ¿El comunicado tiene un resumen?
151. ¿El resumen es demasiado largo o es demasiado breve?
152. ¿El resumen plantea el problema de investigación?
153. ¿El resumen plantea los hallazgos importantes?
154. ¿El resumen es demasiado vago o es muy detallado?

Bibliografía

155. ¿Existe señalamiento bibliográfico de todas las referencias citadas?
156. ¿Es completa cada cita de la bibliografía?
157. ¿Las entradas o señalamientos están ordenados con el mismo estilo bibliográfico, y el estilo es aceptable?

Apéndices

158. ¿Los apéndices se usan eficazmente para presentar materiales largos o suplementarios?
159. ¿Los apéndices están bien organizados y completos?

Estructura y estilo

160. ¿Es satisfactoria la estructura global y el formato del comunicado?
161. ¿El comunicado global muestra detalles suficientes o insuficientes?
162. ¿El estilo de redacción es pomposo y/o pretencioso?
163. ¿El estilo de redacción es demasiado personal y subjetivo?
164. ¿Se evitan implicaciones sexistas del lenguaje?
165. ¿Las frases son demasiado complicadas?
166. ¿Las transiciones son graduales y el comunicado se caracteriza por continuidad de ideas y expresión?
167. ¿Las frases son correctas desde el punto de vista gramatical?
168. ¿Tiene la escritura buena ortografía?
169. ¿Desde el punto de vista estilístico, ¿fue agradable leer el comunicado?

EVALUACIÓN CRÍTICA DE LA INVESTIGACIÓN PUBLICADA *

INTRODUCCIÓN

1. Revisión bibliográfica adecuada
2. Objetivos o hipótesis claramente definidas
3. Selección de una estrategia de investigación apropiada
4. Selección de variables apropiadas

MÉTODOS

5. Modelo de muestreo utilizado
6. Tamaño de la muestra
7. Descripción de la muestra
8. Validez y precisión
9. Descripción de los instrumentos
10. Diseño adecuado
11. Grupos de control
12. Asignación de sujetos
13. Parámetros de tratamiento
14. Efectos Rosenthal y Hawthorne
15. Circunstancias en las que se realiza el estudio
16. Secuencias de tratamientos y observaciones

RESULTADOS

17. Tablas y gráficos adecuados
18. Selección de parámetros estadísticos
19. Cálculo de los parámetros estadísticos

DISCUSIÓN

20. Realización de deducciones correctas a partir de los datos
21. Interpretaciones correctas desde el punto de vista de la lógica
22. Desviaciones inesperadas del protocolo
23. Generalización a partir de los resultados
24. Significación estadística frente a significación clínica
25. Significado teórico del estudio

* Polgar S, Thomas SA. *Introducción a la Investigación en Ciencias de la Salud*. Capítulo 22. Madrid: Chuerchill Livingstone, 1993

LISTA DE COMPROBACIÓN PARA EVALUAR LA INVESTIGACIÓN PUBLICADA *

Problemas que se pueden identificar en un artículo de investigación.	Posibles implicaciones
1. Revisión bibliográfica inadecuada	Definición errónea de la base conceptual de la investigación
2. Objetivos o hipótesis imprecisas	La investigación puede carecer de orientación; la interpretación de la evidencia puede ser ambigua
3. Estrategia de investigación inapropiada	Los resultados pueden no ser relevantes para el problema que se estudia
4. Las variables seleccionadas no son adecuadas	las medidas pueden no estar relacionadas con los conceptos que se investigan
5. Método de muestreo inapropiado	La muestra puede estar sesgada; la investigación carecería de validez externa
6. Tamaño de muestra inapropiado	La muestra puede estar sesgada; el análisis estadístico puede carecer de la potencia necesaria
7. Descripción inadecuada de la muestra	La aplicación de los resultados a grupos o individuos específicos puede ser difícil
8. La instrumentación carece de validez o precisión	Los resultados pueden obedecer a errores de medida
9. Diseño inadecuado	La investigación puede carecer de validez interna; por ejemplo los resultados pueden deberse a variables externas incontroladas
10. Falta de grupos de control adecuados	La investigación puede carecer de validez interna; la magnitud del efecto es difícil de estimar

* Polgar S, Thomas SA. *Introducción a la Investigación en Ciencias de la Salud*. Madrid: Chuerchill Livingstone, 1993

11. Asignación sesgada de sujetos	La investigación puede carecer de validez interna
12. Variaciones o falta de control de los parámetros de tratamiento	La investigación puede carecer de validez interna
13. Los sesgos de los observadores no se controlan (efecto Rosenthal)	La investigación puede carecer de validez interna y externa
14. Las expectativas de los sujetos no se controlan (efectos Hawthorne)	La investigación puede carecer de validez interna y externa
15. Los estudios se realizan en una situación inapropiada	La investigación puede carecer de validez ecológica
16. Falta de control de los tiempos a los que se realizan las observaciones y los tratamientos	Posibles efectos de serie; la investigación puede carecer de validez interna
17. Presentación inadecuada de parámetros estadísticos descriptivos	La naturaleza de los resultados empíricos puede no ser comprensible
18. Parámetros estadísticos inapropiados para describir y/o analizar datos	Distorsión de los datos; pueden derivarse deducciones falsas
19. Cálculo erróneo de parámetros estadísticos	Pueden derivarse deducciones falsas
20. Se efectúan deducciones falsas a partir de los datos	Se pueden establecer conclusiones falsas relativas a los resultados de una investigación
21. Desviaciones del protocolo	La investigación puede carecer de validez externa o interna
22. Generalización errónea de resultados	Puede comprometerse la validez externa
23. Se confunde la significación estadística con la clínica	Pueden favorecerse tratamientos que carecen de utilidad clínica
24. Los resultados carecen de conexión lógica con resultados de estudios previos	Queda sin definir el significado teórico de la investigación

LISTA DE EVALUACION EN LA INVESTIGACION *

COMIENZO

1. ¿La portada indica el título, el nombre de los patrocinadores, investigadores, asesores, universidad, institución, teléfonos, direcciones , fecha de envío?
2. ¿El título es una explicación corta del proyecto?
3. ¿El resumen indica claramente qué se hizo, cómo se llevó a cabo y qué resultados se obtuvieron?
4. ¿La investigación mencionada en el resumen es tan importante que el lector se verá interesado por ésta?
5. ¿Los resultados del trabajo así como el informe aparecen claramente especificados en el resumen?

PROBLEMA

6. ¿La importancia del problema fue transmitida sin exageraciones al lector?
7. ¿El nivel de importancia del problema justificó el presupuesto requerido y/u otros recursos?
8. ¿Si los hallazgos del problema de investigación tuvieron implicaciones más allá de las aplicaciones esperadas, fue ésto mencionado?
9. ¿El objetivo de la investigación fue planteado en términos simples y claros?

REVISION DE LA LITERATURA

10. ¿Los principales hallazgos de la investigación y sus relaciones con el problema fueron indicados?
11. ¿Cuando fue relevante, se indicó la información principal acerca de la organización del estudio y las relaciones de éste?
12. Si se usó una teoría, ¿le queda claro al lector cómo se aplicó ésta a la hipótesis y al problema?
13. ¿Está claro qué aportará el estudio al conocimiento actual?
14. ¿Las referencias fueron citadas apropiadamente?
15. ¿La literatura revisada se encuentra relacionada con la población, las variables o teorías?

* Sproul N. *Handbook of Research Methods*. Chapter 10. London: The Scarecrow Press Inc., 1988

HIPOTESIS O PREGUNTAS DE INVESTIGACION

16. ¿Las hipótesis fueron indicadas cuando hubo la suficiente información para hacer predicciones?
17. ¿Las preguntas de investigación fueron indicadas cuando hubo insuficiente información para la predicción o cuando se deseó exploración de la información periférica hacia el principal problema?
18. ¿Las variables en las hipótesis o preguntas de investigación fueron lo suficientemente claras como para que el lector las haya captado rápidamente?
19. ¿Las relaciones evaluadas en las hipótesis o preguntas de investigación examinadas fueron establecidas adecuadamente?
20. ¿Las hipótesis o preguntas de investigación fueron presentadas como resultados lógicos, surgiendo del problema de investigación y de la revisión de la información y de la literatura?
21. ¿Fueron indicadas las hipótesis de investigación más que las hipótesis de nulidad?
22. ¿Las preguntas o hipótesis de investigación fueron indicadas en la forma en que se debería de probar?
23. ¿Cada variable es definida operacionalmente en términos medibles?

POBLACION Y MUESTRA

24. ¿La población fue claramente definida en su descripción?
25. ¿El tamaño de la población fue dado?
26. ¿El método de muestreo fue especificado completamente?
27. Si se utilizó estratificación, ¿existe un informe acerca de por qué se utilizó y cómo se conformaron los estratos?
28. ¿Existe algún informe que justifique el procedimiento utilizado si la muestra no incluyó procedimientos aleatorios?
29. ¿Fue adecuado el tamaño de la muestra y los criterios para su selección?

DISEÑO

30. ¿Fue claro el tipo de diseño de investigación utilizado?
31. Si se utilizó diseño experimental o cuasiexperimental, ¿fue descrito y explicado?
32. Si se utilizó la estructura experimental, ¿los diferentes niveles de tratamiento son especificados y explicados?
33. Si se utilizó la estructura experimental, ¿se mencionó si los sujetos fueron asignados aleatoriamente al tratamiento?
34. Si se utilizó la estructura cuasiexperimental, ¿hay alguna explicación de por qué la estructura experimental no pudo utilizarse?
35. ¿Han sido utilizados procedimientos para disminuir problemas de validez interna y externa?
36. ¿Fue impuesto al diseño tanto control como el factible?
37. Si se utilizó una estructura de investigación histórica, ¿se da suficiente información acerca de la representatividad y posibles distorsiones u omisiones en los datos obtenidos de fuentes primarias o secundarias?

METODOS DE OBTENCION DE DATOS

38. ¿El método de obtención de datos es el más adecuado y factible para valorar la hipótesis o preguntas de investigación?
39. ¿Podría una recolección de datos menos costosa proveer buena información?
40. Cuando se entrevista, administran instrumentos o se observa, ¿Se dedica tiempo y recursos al entrenamiento?
41. Si se examinaron documentos, materiales o artefactos, ¿los resultados son lo suficientemente representativos y confiables para trazar conclusiones?
42. Si se utilizaron observaciones o evaluación de documentos o materiales, ¿hubo muestreo de evento o tiempo, o se realizó con segmentos totales?
43. ¿El método de recolección de datos afectó la respuesta de los sujetos? Si fue así, ¿el efecto es lo suficientemente serio para cambiar los métodos de recolección de datos?

INSTRUMENTACION

44. ¿Mide las variables apropiadamente el instrumento?
45. ¿Los niveles de medición proporcionados son los mismos que el investigador planeó?
46. ¿Tiene el instrumento evidencia de validez, en cuanto a magnitud y tipo?
47. ¿Tiene el instrumento evidencia de seriedad, en cuanto a magnitud y tipo?
48. ¿Pueden los sujetos responder al instrumento en un período razonable de tiempo?
49. ¿Es lo suficientemente sencillo el instrumento para los sujetos que lo van a responder?
50. ¿Es el instrumento fácil de administrar en un período de tiempo razonable?
51. ¿Los resultados del instrumento son fáciles de interpretar?
52. ¿Tendrá el instrumento efectos psicológicos o sociales? Si es así, ¿qué medidas se incorporan para manejar esos efectos?
53. Si el instrumento es nuevo o una versión modificada, ¿hubo una prueba piloto?
54. ¿Hay una copia de cada instrumento en los apéndices?

PRUEBAS ESTADISTICAS Y ANALISIS

55. ¿Son las pruebas estadísticas apropiadas para probar la hipótesis o responder la(s) pregunta(s) de investigación?
56. ¿Los niveles de medición de las variables coinciden con los niveles requeridos para los procedimientos estadísticos?
57. Si las observaciones se relacionan entre sí, ¿la prueba estadística es la apropiada para respuestas no independientes?
58. ¿Existen reservas de los datos examinados para determinar cualquier resultado no anticipado?
59. ¿Están incluidas las medidas descriptivas y de comprensión de datos (como distribución de frecuencias, medidas de tendencia central, medidas de variabilidad, tablas cruzadas, entre otras), para aumentar la capacidad de interpretación de los resultados?

CRONOGRAMA

60. ¿Se describen de una forma sencilla las fases de la investigación?
61. ¿El tiempo estimado fue consistente con los recursos personales descritos?

PERSONAL

62. ¿Se mencionan por nombre y área el director del proyecto y los asociados principales?

REFERENCIAS

63. ¿Se dieron las referencias de cada cita en el trabajo?
64. ¿Son apropiadas las referencias para los conceptos descritos?
65. ¿Tienen balance las referencias entre los autores clásicos y los más recientes?
66. ¿Se distingue a las referencias citadas de la información bibliográfica?
67. ¿Las referencias están correctamente colocadas en el trabajo y en los apéndices?

APENDICES

68. ¿Aparecen en los apéndices, datos, tablas, esquemas y otras informaciones extensas de modo que la lectura del trabajo no se vea interrumpida?
69. ¿Los apéndices están claramente delineados por títulos, notas y números?

GENERAL

70. Si hay consideraciones éticas ¿hay alguna explicación de cómo se manejaron?
71. ¿El trabajo tiene errores gramaticales o numéricos?
72. ¿Puede leer el trabajo un no especialista y entender los objetivos?
73. ¿Es el formato del trabajo lo suficientemente sucinto como para leerlo fácil y rápidamente sin perder los mayores puntos?

VARIOS

74. Mandar el trabajo a tiempo. Generalmente los comités rechazan los trabajos que llegan después de la fecha límite.
75. Si el trabajo se manda por correo, de preferencia que éste sea de primera clase y certificado.
76. La dirección y destino deben aparecer claramente (incluyendo de ser posible el nombre del representante de la comisión científica).
77. Evitar que la apariencia física del trabajo refleje extravagancia. Algunos investigadores invierten tiempo en ello, esperando impresionar a los revisores. Esto frecuentemente tiene un efecto opuesto.

LISTA PARA EVALUACION DE ARTICULOS MEDICOS GENERALES *

DISEÑO DEL ESTUDIO	SI	DUDOSO	NO
1. ¿El objetivo del estudio se describe suficientemente?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ¿El diseño del estudio se describe suficientemente?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ¿El diseño del estudio es apropiado y objetivo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ¿La fuente de los sujetos se describe claramente?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. ¿El método de selección de los sujetos se describe claramente? (criterios de inclusión y de exclusión)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. ¿La muestra es apropiada con respecto a la población a la cual se referirán los hallazgos?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. ¿El tamaño de la muestra se basó en consideraciones preestudio de poder estadístico?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. ¿El diseño del estudio es aceptable?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
 CONDUCCION DEL ESTUDIO			
9. ¿Se logró una tasa satisfactoria de respuestas?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
 ANALISIS Y PRESENTACION			
10. ¿Hay una declaración que describa adecuadamente o que dé referencias de todos los procedimientos estadísticos usados?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. ¿Los métodos estadísticos usados son apropiados para los datos?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. ¿Se los usa correctamente?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

* Altman DG. *Practical Statistics for Medical Research*.. London: Chapman and Hall, 1991

13. La presentación de material estadístico (tablas, gráficos, etc.) ¿es satisfactoria?
14. ¿Se presenta suficiente análisis?
15. ¿Se dan intervalos de confianza para los resultados principales?

EVALUACION GLOBAL

16. ¿Se justifican las conclusiones sacadas del análisis estadístico?
17. ¿El artículo es aceptable estadísticamente?

LISTA PARA EVALUACION DE INFORMES DE ENSAYOS CLINICOS *

DISEÑO DEL ESTUDIO	SI	DUDOSO	NO
1. ¿Se describe suficientemente el objetivo del ensayo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ¿Se describe suficientemente el diseño del estudio?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ¿Hay afirmaciones satisfactorias de los criterios diagnósticos para inclusión en el ensayo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ¿Se describe claramente la fuente de los sujetos?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. ¿Se definen bien los tratamientos?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. ¿Se estudian convenientemente los grupos de tratamiento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. ¿Se utiliza la asignación aleatoria al tratamiento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. ¿Se describe el método de creación de la aleatorización (Ej. tablas de números aleatorios)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. ¿Se describe el mecanismo de asignación al tratamiento (sobres sellados)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. ¿Se diseña el mecanismo de asignación al tratamiento para eliminar sesgos?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. ¿Hay un retraso aceptablemente corto desde el momento de la asignación hasta el inicio del tratamiento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. ¿Se utiliza el grado potencial de «ceguera» durante el ensayo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. ¿Hay una afirmación satisfactoria de criterios para medidas de impacto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. ¿Son adecuadas las medidas de impacto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. ¿Hay una descripción de un cálculo de preestudio del tamaño de la muestra basado en consideraciones de poder estadístico?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

* Altman DG. *Practical Statistics for Medical Research*.. London: Chapman and Hall, 1991

16. ¿Se especifica el seguimiento de la duración del posttratamiento?
17. ¿Es aceptable el diseño del estudio?

CONDUCCION DEL ESTUDIO

18. ¿Se hace seguimiento a una alta proporción de sujetos?
19. ¿La proporción de sujetos que completan el estudio es alta?
20. ¿Los abandonos (deserciones) se describen separadamente para cada grupo de tratamiento?
21. ¿Los efectos laterales del tratamiento se describen separadamente para cada grupo?

ANALISIS Y PRESENTACION

22. ¿Hay una afirmación que describe adecuadamente o se citan las referencias de todos los procedimientos estadísticos utilizados?
23. ¿Se presenta adecuadamente la línea de base de las características de cada grupo?
24. ¿Los métodos estadísticos utilizados son apropiados para los datos?
25. ¿Estos métodos se utilizan correctamente?
26. ¿Se consideran adecuadamente los factores pronósticos?
27. ¿Es satisfactoria la presentación del material estadístico? (tablas, gráficos, numérico).
28. ¿Se presenta un análisis suficiente?
29. ¿Se dan los intervalos de confianza para los resultados principales?

EVALUACION GLOBAL

30. ¿Se justifican las conclusiones sacadas del análisis estadístico?
31. ¿El artículo es estadísticamente aceptable?

APÉNDICE IV

REQUISITOS UNIFORMES PARA PREPARAR LOS MANUSCRITOS ENVIADOS A REVISTAS BIOMÉDICAS ¹

(5ª Edición, 1998)

Comité Internacional de Editores de Revistas Médicas ²

En 1978, un reducido grupo de directores de revistas médicas generales que se publican en inglés se reunió de manera informal en Vancouver (Canadá) a fin de fijar normas con respecto al formato que deberían adoptar los manuscritos enviados a esas publicaciones. Este fue el inicio de lo que con el tiempo llegó a conocerse como el Grupo de Vancouver. Sus requisitos para la preparación de manuscritos, que incluían el formato de las referencias bibliográficas creado por la Biblioteca Nacional de Medicina de los Estados Unidos, se publicaron por vez primera en 1979. Con el paso del tiempo el Grupo de Vancouver creció y se convirtió en el Comité Internacional de Directores de Revistas Médicas (CIDRM), que se reúne una vez al año y que gradualmente ha venido ampliando los temas que le conciernen.

El Comité ha producido cinco ediciones de los requisitos uniformes para preparar los manuscritos enviados a revistas biomédicas. A lo largo de los años han surgido asuntos que van más allá de la preparación del manuscrito. Algunos de esos asuntos se han incorporado a los requisitos uniformes, mientras que otros se tratan en declaraciones por separado. Cada declaración se ha publicado en una revista científica.

La quinta edición (1997) de los requisitos uniformes se ha preparado con la finalidad de reorganizar y modificar la redacción de la cuarta edición a efecto de aumentar la claridad y abordar inquietudes con respecto a derechos, protección de la vida privada de los sujetos de inves-

¹ Traducción al español basada en: International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to bio-medical journals. *Ann Intern Med* 1997; 126:36-47. La presente traducción y la de las declaraciones que el Comité Internacional de Directores de Revistas Médicas publica por separado se podrán encontrar en el siguiente sitio de la World Wide Web: <http://www.paho.org/spanish/DBI/authors.htm>

Traductor: Dr. Gustavo A. Silva, OPS, Washington, DC, EUA. Dirección postal: Organización Panamericana de la Salud, AGS/T, 525 Twentythird Street, Washington, DC 20037, EUA. Correo electrónico: lsilvagu@paho.org

El documento original en inglés se puede consultar en versión digital en diversos sitios de la World Wide Web, entre otros ACP Online (<http://www.acponline.org>).

Este documento no está protegido por derechos de autor. Puede copiarse o reimprimirse sin autorización, siempre y cuando se haga sin fines de lucro.

Las consultas y observaciones deben dirigirse a Kathleen Case, ICMJE Secretariat Office, *Annals of Internal Medicine*, American College of Physicians Independence Mall West, Sixth Street at Race, Philadelphia, PA 19106 1572, EUA. Teléfono: 215-351-2661; fax: 215-351-2644; correo electrónico: Kathyc@acp.mhs.compuserve.com

² Actualmente integran el comité editores afiliados con las siguientes revistas y entidades: *Annals of Internal Medicine*, Biblioteca Nacional de Medicina de los Estados Unidos, *British Medical Journal*, *Canadian Medical Association Journal*, *The Journal of the American Medical Association*, *The Lancet*, *The Medical Journal of Australia*, *The New England Journal of Medicine*, *The New Zealand Medical Journal*, *Tidsskrift for den Norske Lægeforening*, Universidad de Princeton y *Western Journal of Medicine*.

tigación, descripción de los métodos aplicados y otros asuntos. Los requisitos uniformes para preparar los manuscritos enviados a revistas biomédicas pueden reproducirse para fines educativos y sin afán de lucro, con prescindencia de los derechos de autor; el Comité alienta la distribución de este material.

A las revistas que accedan a guiarse por los requisitos uniformes (actualmente más de 500 revistas lo hacen) se les pide que en sus instrucciones para los autores citen el documento correspondiente a 1997.

Es importante hacer hincapié en lo que los requisitos significan y en lo que no. En primer lugar, los requisitos uniformes son instrucciones sobre la forma de preparar los manuscritos dirigidas a los autores; no se trata de indicaciones sobre estilo editorial destinadas a la redacción de las revistas. (Sin embargo, muchas revistas han incorporado en su estilo editorial ciertos elementos de los requisitos.)

En segundo lugar, si los autores preparan sus manuscritos según el estilo especificado en estos requisitos, los directores de las revistas participantes no devolverán los manuscritos para que se hagan cambios de estilo antes de considerarlos para publicación. No obstante, en el proceso de publicación los manuscritos podrán ser modificados por las redacciones de las revistas para adaptarlos a las particularidades de su propio estilo editorial.

En tercer lugar, los autores que envíen manuscritos a una revista participante se abstendrán de prepararlos de acuerdo con el estilo editorial de ésta y se concretarán a cumplir con los requisitos uniformes.

Pese a todo, los autores tendrán asimismo que seguir las instrucciones particulares de la revista en cuestión por lo que se refiere a los temas apropiados para ésta y el tipo de manuscritos que se le pueden enviar: por ejemplo, artículos originales, revisiones o informes de casos. Además, es probable que en dichas instrucciones figuren otros requisitos exclusivos de la publicación, tales como el número de copias del manuscrito, los idiomas en que éste puede ser redactado, la extensión de los artículos y las abreviaturas aprobadas.

Se espera que las revistas participantes declaren en sus instrucciones a los autores que sus normas están de acuerdo con los requisitos uniformes para preparar los manuscritos enviados a revistas biomédicas y citen una versión publicada de éstos.

CUESTIONES QUE DEBEN CONSIDERARSE ANTES DE PRESENTAR UN MANUSCRITO PARA PUBLICACION

Publicación redundante o duplicada

Por publicación redundante o duplicada se entiende la publicación de un artículo que se traslapa considerablemente con otro ya publicado.

Los lectores de publicaciones periódicas primarias merecen que se les dé la confianza de que lo que están leyendo es original, a menos que haya una clara indicación de que el artículo se ha vuelto a publicar por elección del autor y el director. Esta posición tiene como fun-

damento las leyes internacionales de derechos de autor, la conducta ética y el uso eficaz en función de los costos de los recursos.

La mayor parte de las revistas científicas no desean recibir manuscritos acerca de un trabajo que ya se ha dado a conocer en gran medida en un artículo publicado o que forma parte de otro manuscrito que se ha propuesto o ha sido aceptado para publicación en otra parte, ya sea en forma impresa o en soporte electrónico. Esta norma no impide que la revista considere un artículo rechazado por otra revista o un informe completo que sigue a la publicación de un informe preliminar, como puede ser un resumen o póster presentados a la consideración de colegas en una reunión profesional. Tampoco impide que las revistas consideren un artículo que se ha presentado en una reunión científica pero que no se ha publicado íntegramente ni se está considerando para publicación en las actas de una reunión o en una publicación semejante. Las informaciones periodísticas acerca de reuniones programadas no se considerarán en general como infracciones de esta regla, pero no habrán de ampliarse mediante datos suplementarios o copias de los cuadros y las ilustraciones.

Cuando se envíe un artículo para publicación, el autor debe siempre adjuntar una relación completa de toda presentación del documento a otras revistas y de cualquier informe anterior que pudieran considerarse publicación redundante o duplicada del mismo trabajo o de uno muy semejante. El autor debe poner sobre aviso al director de la revista si el trabajo aborda temas sobre los cuales se hayan publicado informes anteriores. Si tal es el caso, esos trabajos se mencionarán con la debida referencia bibliográfica en el artículo nuevo. Junto con el manuscrito propuesto se incluirán copias de dichos materiales para ayudar al director a decidir la manera de manejar este asunto.

Si la publicación redundante o duplicada se intenta o se produce sin que medie la notificación descrita, los autores deberán atenerse a las medidas editoriales que se tomen en su contra. Como mínimo, cabe esperar el pronto rechazo del manuscrito presentado. Si el director no estaba al tanto de las infracciones y el artículo acaba por aparecer en su revista, entonces probablemente se publique en ésta un aviso de publicación redundante o duplicada, que puede o no acompañarse de una explicación del autor y no necesitará de su aprobación.

La divulgación preliminar, generalmente por conducto de los medios de comunicación de masas, de la información científica descrita en un artículo ya aceptado pero aún sin publicar representa una infracción de las normas de muchas revistas. En contadas ocasiones, y solo mediante acuerdo con el director, puede aceptarse la diseminación preliminar de datos; por ejemplo, cuando se presenta una emergencia de salud pública.

Publicación secundaria aceptable

La publicación secundaria en el mismo idioma o en otro distinto, especialmente en otros países, se considera justificable y puede incluso ser beneficiosa, siempre y cuando se cumplan las siguientes condiciones:

1. Los autores tendrán que recabar la aprobación de los directores de las dos revistas involucradas; el director de la publicación secundaria tendrá en su poder una fotocopia, separata o manuscrito de la versión primaria.

2. Se respetará la precedencia de la publicación primaria dejando transcurrir por lo menos una semana antes de la publicación secundaria (a menos que específicamente se negocie otra cosa con ambos directores).
3. El artículo para publicación secundaria se destinará a un grupo diferente de lectores; podría bastar con una versión abreviada.
4. La versión secundaria reflejará fielmente los datos y las interpretaciones de la primaria.
5. Mediante una nota colocada al pie de la primera página de la versión secundaria, se informará a los lectores, a los colegas de los autores y a los organismos de documentación que el artículo ya se ha publicado con anterioridad en forma total o parcial, indicando la referencia primaria. Este podría ser un texto apropiado para dicha nota: “El presente artículo está basado en un estudio que se dio a conocer primero en [título de la revista y referencia completa]” .

El permiso para la publicación secundaria de este tipo se otorgará gratuitamente.

Protección de los derechos del paciente a que se respete su vida privada

El derecho de los pacientes a que se respete su vida privada no debe ser infringido sin antes obtener su consentimiento fundamentado. Las descripciones por escrito, las fotografías y los árboles genealógicos que se publiquen no deberán contener información por la cual se pueda identificar a los pacientes, a menos que dichos datos resulten esenciales para fines científicos y que el paciente (o su padre o tutor) otorgue por escrito su autorización para la publicación después de haber sido debidamente informado. Para obtener el consentimiento para esta finalidad, es preciso mostrarle al paciente el manuscrito que se va a publicar.

Los detalles que puedan revelar la identidad del paciente deben omitirse si no son esenciales, pero los datos del paciente nunca se alterarán ni falsificarán solamente por tratar de lograr el anonimato. Es difícil lograr el anonimato completo, y frente a cualquier duda será preciso obtener el consentimiento fundamentado. Por ejemplo, cubrir la región de los ojos en las fotografías de los pacientes es una protección insuficiente del anonimato.

La exigencia del consentimiento con conocimiento de causa debe figurar en las instrucciones para los autores de la revista. Siempre que se obtenga dicha anuencia, esto deberá constar en el artículo publicado.

REQUISITOS PARA LA PRESENTACION DE MANUSCRITOS A UNA REVISTA

Resumen de los requisitos técnicos

- Todas las partes del manuscrito estarán a doble espacio.
- Cada sección o componente comenzará en página nueva.
- Revise la secuencia: página del título, resumen y palabras clave, texto, agradecimientos, referencias, cuadros (cada uno en página aparte), pies o epígrafes de las ilustraciones.
- Las ilustraciones se presentarán en forma de impresiones fotográficas sin montar, y no deberán exceder de 203 x 254 mm.
- Incluya la autorización para reproducir material publicado con anterioridad o para usar ilustraciones en las que se pueda identificar a los sujetos humanos.

- Adjunte la transferencia de los derechos de autor y otros formularios.
- Presente el número exigido de copias impresas del artículo.
- Guarde copias de todo lo que envíe.

Preparación del manuscrito

El texto de los artículos de observación y experimentales se divide generalmente, aunque no por fuerza, en secciones que llevan estos encabezamientos: introducción, métodos, resultados y discusión. En los artículos largos puede ser necesario agregar subtítulos dentro de estas secciones, sobre todo en las de resultados y discusión, a fin de hacer más claro el contenido. Es probable que otro tipo de artículos -como los informes de casos, las revisiones y los editoriales- exijan otra estructura. Para mayor orientación, los autores deberán consultar la revista en la que pretenden publicar.

Mecanografíese el manuscrito en papel bond blanco de 216 x 279 mm o de la medida estándar ISO A4 (212 x 297 mm), con márgenes de por lo menos 25 mm. Escríbase o imprímase solamente sobre una cara del papel. Utilícese doble espacio a lo largo de todo el manuscrito, incluidos la página del título, el resumen, el texto, los agradecimientos, las referencias, cada uno de los cuadros y los pies o epígrafes de las ilustraciones. Numérense las páginas en forma consecutiva, empezando por la del título. Sobre el ángulo superior o inferior derecho de cada página anótese el número que le corresponde.

Manuscritos en disquete

Cuando un artículo se halla cercano a la aceptación definitiva, algunas revistas piden que los autores faciliten una copia en forma electrónica (en disquete); pueden aceptar una variedad de formatos de procesadoras de texto o ficheros de texto (ASCII).

Cuando presenten disquetes, los autores deberán:

- 1) cerciorarse de incluir la versión impresa del artículo que va en el disquete;
- 2) poner en el disquete únicamente la versión más reciente del manuscrito;
- 3) denominar claramente el fichero;
- 4) rotular el disquete con indicación del formato y el nombre del fichero;
- 5) facilitar información sobre el equipo de computación y el software utilizados.

Los autores deberán consultar las instrucciones que la revista proporciona a los autores para determinar cuáles son los formatos que se aceptan, las convenciones que se aplican para denominar los ficheros, el número de copias que deben presentarse y otros detalles.

Página del título

La primera página contendrá: 1) el título del artículo, que será conciso pero informativo; 2) nombre de pila preferido y apellidos de cada autor, acompañados de sus grados académicos más importantes y su afiliación institucional; 3) nombre del departamento o departamentos y la institución o instituciones a los que se debe atribuir el trabajo; 4) declaraciones de descargo de responsabilidad, si las hay; 5) nombre y dirección del autor que se ocupará de la correspondencia relativa al manuscrito; 6) nombre y dirección del autor a quien se dirigirá

las solicitudes de separatas, o nota informativa de que los autores no las proporcionarán; 7) procedencia del apoyo recibido en forma de subvenciones, equipo, medicamentos o todo ello; y 8) título abreviado (titulillo) que no pase de 40 pulsaciones (contando caracteres y espacios), el cual se colocará, debidamente identificado como tal, en la última línea de la página inicial.

Autoría

Todas las personas designadas como autores habrán de cumplir con ciertos requisitos para tener derecho a la autoría. Cada autor debe haber participado en el trabajo en grado suficiente para asumir responsabilidad pública por su contenido.

Para concederle a alguien el crédito de autor, hay que basarse únicamente en su contribución esencial por lo que se refiere a los siguientes aspectos: 1) la concepción y el diseño o bien el análisis y la interpretación de los datos; 2) la redacción del artículo o la revisión crítica de una parte importante de su contenido intelectual; y 3) la aprobación final de la versión que será publicada. Las tres condiciones tendrán que cumplirse siempre. La participación que consiste meramente en conseguir financiamiento o recoger datos no justifica el crédito de autor. Tampoco basta con ejercer la supervisión general del grupo de investigación. Toda parte del artículo que sea decisiva con respecto a las conclusiones principales deberá ser responsabilidad de por lo menos uno de los autores.

Los directores de revista podrán solicitar a los autores que describan la contribución de cada uno; esa información puede ser publicada.

Cada vez es más común que los ensayos multicéntricos se atribuyan a un autor corporativo. Todos los miembros del grupo que sean designados como autores, ya sea en la línea destinada al nombre de los autores a continuación del título o en una nota a pie de página, deberán cumplir plenamente con los requisitos de autoría recién señalados. Los miembros del grupo que no cumplan con dichos criterios serán mencionados, con su autorización, en la sección de agradecimientos o en un apéndice (véase "Agradecimientos").

El orden en que figuran los autores debe reflejar una decisión conjunta de estos. Como los autores se suelen enumerar de distintas maneras, el significado del orden en que aparecen no puede deducirse con exactitud a menos que ellos mismos lo enuncien explícitamente. Para tal efecto, tal vez deseen agregar, en una nota a pie de página, la explicación sobre el orden de enumeración. Al decidir acerca de dicho orden, los autores tendrán presente que muchas revistas imponen un límite al número de autores que figuran en el índice de materias y que, cuando hay más de 25 autores, la Biblioteca Nacional de Medicina de los Estados Unidos incluye en MEDLINE tan sólo los nombres de los 24 primeros mas el del último.

Resumen y palabras clave

La segunda página incluirá un resumen (que no sobrepasará las 150 palabras de extensión si es un resumen ordinario o las 250 si es uno estructurado). En él se indicarán los propósitos del estudio o investigación; los procedimientos básicos (selección de los sujetos o los animales de laboratorio incluidos en el estudio; métodos de observación y análisis); los hallazgos más importantes (proporcionense datos específicos y, de ser posible, su significación esta-

dística); y las conclusiones principales. Hágase hincapié en los aspectos nuevos e importantes del estudio o las observaciones.

A continuación del resumen agréguese, debidamente rotuladas, de 3 a 10 palabras o frases cortas clave que ayuden a los indizadores a clasificar el artículo, las cuales se publicarán junto con el resumen. Utilícense para este propósito los términos de la lista “Medical Subject Headings” (MeSH) [Encabezamientos de temas médicos] del *Index Medicus*; en el caso de términos de reciente aparición que todavía no figuren en dicha lista, podrán usarse las expresiones corrientes.

Introducción

Expresa el propósito del artículo y resume el fundamento lógico del estudio u observación. Mencione las referencias estrictamente pertinentes y no incluya datos ni conclusiones del trabajo que está dando a conocer.

Métodos

Describe claramente la forma como se seleccionaron los sujetos observados o que participaron en los experimentos (pacientes o animales de laboratorio, incluidos los testigos).

Identifique la edad, el sexo y otras características importantes de los sujetos. La definición y la pertinencia de la raza y el grupo étnico son ambiguos. Los autores deberán ser particularmente cuidadosos con respecto a usar estas categorías.

Identifique los métodos, los aparatos (nombre y dirección del fabricante entre paréntesis) y los procedimientos con detalles suficientes para que otros investigadores puedan reproducir los resultados. Proporcione referencias de los métodos acreditados, incluidos los de índole estadística (véase más adelante); dé referencias y explique brevemente los métodos ya publicados pero que no son bien conocidos; describa los métodos nuevos o que han sido sustancialmente modificados, manifestando las razones por las cuales se usaron y evaluando sus limitaciones. Identifique exactamente todos los medicamentos y productos químicos utilizados, sin olvidar nombres genéricos, dosis y vías de administración.

Los informes de ensayos clínicos aleatorizados deberán presentar información sobre todos los elementos importantes del estudio, como son el protocolo (población de estudio, intervenciones o exposiciones, resultados y el fundamento lógico del análisis estadístico), asignación de intervenciones (métodos de aleatorización, ocultamiento de la asignación a los grupos de tratamiento) y método de enmascaramiento (método ciego).

Los autores que presenten manuscritos de revisión incluirán una sección en la que se describan los métodos utilizados para localizar, seleccionar, extraer y sintetizar los datos. Estos métodos se mencionarán también en forma sinóptica en el resumen.

Ética. Cuando informe sobre experimentos en seres humanos, señale si los procedimientos seguidos estuvieron de acuerdo con las normas éticas del comité (institucional o regional) que supervisa la experimentación en seres humanos o con la Declaración de Helsinki de 1975, modificada en 1983. No utilice el nombre de los pacientes, sus iniciales ni los códigos hos-

pitalarios, especialmente en el material ilustrativo. Cuando dé a conocer experimentos con animales, mencione si se cumplieron las normas de la institución, las de un consejo nacional de investigación o cualquier ley nacional acerca del cuidado y el uso de animales de laboratorio.

Estadística. Describa los métodos estadísticos con detalles suficientes para que el lector versado en el tema y que tenga acceso a los datos originales pueda verificar los resultados presentados. Siempre que sea posible, cuantifique los resultados y preséntelos con indicadores apropiados de error o incertidumbre de la medición (por ej., intervalos de confianza). No dependa exclusivamente de las pruebas de comprobación de hipótesis estadísticas, tales como el uso de los valores P , que no transmiten información cuantitativa importante. Analice la elegibilidad de los sujetos de experimentación. Proporcione los detalles del proceso de aleatorización. Describa los medios utilizados para enmascarar las observaciones (método ciego), indicando los resultados que dieron. Informe sobre las complicaciones del tratamiento. Especifique el número de observaciones. Mencione las pérdidas de sujetos de observación (por ej., las personas que abandonan un ensayo clínico). Siempre que sea posible, las referencias sobre el diseño del estudio y los métodos estadísticos utilizados serán de trabajos vigentes (indicando el número de las páginas), y no de los artículos originales donde se describieron por vez primera. Especifique cualquier programa de computación de uso general que se haya empleado.

Las descripciones generales de los métodos utilizados deben aparecer en la sección de métodos. Cuando resuma los datos en la sección de resultados, especifique los métodos estadísticos que se emplearon para analizarlos. Limite el número de cuadros y figuras al mínimo necesario para explicar el tema central del artículo y para evaluar los datos en que se apoya. Use gráficas en vez de cuadros subdivididos en muchas partes; no duplique los datos en las gráficas y los cuadros. Evite el uso no técnico de términos de la estadística, tales como “al azar” (que entraña el empleo de un método de aleatorización), “normal”, “significativo”, “correlaciones” y “muestra”. Defina los términos, las abreviaturas y la mayor parte de los símbolos estadísticos.

Resultados

En el texto, los cuadros y las ilustraciones, presente los resultados siguiendo una secuencia lógica. No repita en el texto todos los datos de los cuadros ni de las ilustraciones; destaque o resuma tan sólo las observaciones importantes.

Discusión

Haga hincapié en los aspectos nuevos e importantes del estudio y en las conclusiones que se derivan de ellos. No repita con pormenores los datos u otra información ya presentados en las secciones de introducción y de resultados. Explique en la sección de discusión el significado de los hallazgos y sus limitaciones, incluidas sus implicaciones para la investigación futura. Relacione las observaciones con otros estudios pertinentes.

Establezca el nexo entre las conclusiones y los objetivos del estudio, pero absténgase de hacer afirmaciones generales y extraer conclusiones que no estén completamente respaldadas por los datos. En particular, los autores evitarán hacer afirmaciones sobre los beneficios y los

costos económicos, a menos que su manuscrito incluya datos y análisis económicos. No reclame ningún tipo de precedencia ni mencione trabajos que no estén terminados. Proponga nuevas hipótesis cuando haya justificación para ello, pero identificándolas claramente como tales. Cuando sea apropiado, puede incluir recomendaciones.

Agradecimientos

En un lugar adecuado del artículo (como nota al pie de la primera página o como apéndice del texto; véanse los requisitos de la revista) uno o varios enunciados especificarán lo siguiente: 1) las colaboraciones que deben ser reconocidas pero que no justifican la autoría, tales como el apoyo general del jefe del departamento; 2) el reconocimiento por la ayuda técnica recibida; 3) el agradecimiento por el apoyo financiero y material, especificando la índole del mismo; y 4) las relaciones que puedan suscitar un conflicto de intereses (véase “Conflicto de intereses”).

Las personas que colaboraron intelectualmente en el artículo pero cuya participación no justifica la autoría pueden ser citadas por su nombre, añadiendo su función o tipo de colaboración; por ejemplo, “asesoramiento científico”, “examen crítico de la propuesta para el estudio”, “recolección de los datos” o “participación en el ensayo clínico”. Estas personas tendrán que conceder su permiso para ser nombradas. Los autores se responsabilizarán de obtener la autorización por escrito de las personas mencionadas por su nombre en los agradecimientos, pues los lectores pueden inferir que éstas respaldan los datos y las conclusiones.

El reconocimiento por la ayuda técnica recibida figurará en un párrafo separado de los testimonios de gratitud por otras contribuciones.

Referencias

Numere las referencias consecutivamente siguiendo el orden en que se mencionan por primera vez en el texto. En éste, en los cuadros y en los pies o epígrafes de las ilustraciones, las referencias se identificarán mediante números arábigos entre paréntesis. Las referencias citadas solamente en cuadros o ilustraciones se numerarán siguiendo una secuencia que se establecerá por la primera mención que se haga en el texto de ese cuadro o esa figura en particular.

Emplee el estilo de los ejemplos que aparecen más adelante, los cuales están basados en el formato que la Biblioteca Nacional de Medicina de los Estados Unidos usa en el *Index Medicus*. Abrevie los títulos de las revistas de conformidad con el estilo utilizado en dicha publicación. Consulte la *List of Journals Indexed in Index Medicus* [Lista de revistas indexadas en *Index Medicus*], que se publica anualmente como parte del número de enero y como separata. La lista se puede obtener asimismo en el sitio que la biblioteca mantiene en la World Wide Web (<http://www.nlm.nih.gov>).

Absténgase de utilizar los resúmenes como referencias. Las referencias a artículos que han sido aceptados pero que todavía no se publican se designarán como “en prensa” o “de próxima aparición”; los autores obtendrán por escrito el permiso para citar dichos artículos y también la verificación de que han sido aceptados para publicación. La información prove-

niente de manuscritos presentados para publicación pero aún no aceptados se citará en el texto como “observaciones inéditas”, con el permiso correspondiente de la fuente.

No cite una “comunicación personal” a menos que aporte información esencial que no pueda obtenerse de una fuente pública; en ese caso, el nombre de la persona y la fecha de la comunicación aparecerán entre paréntesis en el texto. En el caso de artículos científicos, los autores deberán obtener el permiso de la fuente y su confirmación de la exactitud de la comunicación personal, ambos por escrito.

Los autores verificarán las referencias cotejándolas contra los documentos originales.

El estilo de los requisitos uniformes (estilo de Vancouver) se basa en gran medida en una norma de estilo ANSI adaptada por la Biblioteca Nacional de Medicina (NLM) para sus bases de datos. En los ejemplos que siguen se han agregado notas cuando el estilo de Vancouver difiere del estilo que actualmente utiliza la NLM.

Artículos de revista

1. Artículo de revista ordinario

Enumere los primeros seis autores y añada la expresión “et al.”

(Nota: La NLM incluye ahora hasta 25 autores; si hay más de 25, enumera los primeros 24, a continuación el último autor y luego agrega “et al.”)

Vega KI, Pina I, Krevsky B Heart transplantation is associated with an increased risk for pancreatobiliary disease *Ann Intern Med* 1996 Jun 1;124(11):980-3.

Optativamente, si se utiliza la paginación continua a lo largo de un volumen (como hacen muchas revistas médicas), se pueden omitir el mes y el número.

(Nota: Para respetar la uniformidad, en todos los ejemplos que se presentan en los requisitos uniformes se aplica esta opción. La NLM, sin embargo, no usa dicha opción.)

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. *Ann Intem Med* 1996;124:980-3.

Más de seis autores:

Parkin DM, Clayton D, Black RJ, Masuyer E, Friedl HP, Ivanov E, et al. Childhood leukaemia in Europe after Chernobyl: 5 year follow-up. *Br J Cancer* 1996;73:1006-12.

2. Organización como autor

The Cardiac Society of Australia and New Zealand. Clinical exercise stress testing. Safety and performance guidelines. *Med J Aust* 1996;164:282-4.

3. No se indica el nombre del autor

Cancer in South Africa [editorial]. *S Afr Med* 1 1994;84:15

4. *Artículo en idioma extranjero*³

(Nota: La NLM traduce el título al inglés, lo encierra entre corchetes y le agrega la abreviatura correspondiente al idioma original.)

Ryder TE, Haukeland EA, Solhaug JH. Bilateral infrapatellar seneruptur hos tidligere frisk kvinne. Tidsskr Nor Laegeforen 1996; 116:41-2.

5. *Suplemento de un volumen*

Shen HM, Zhang QF. Risk assessment of nickel carcinogenicity and occupational lung cancer. Environ Health Perspect 1994;102 Suppl 1:275-82.

6. *Suplemento de un número*

Payne DK, Sullivan MD, Massie MJ. Women's psychological reactions to breast cancer Semin Oncol 1996;23(1 Suppl 2): 89-97.

7. *Parte de un volumen*

Ozben T, Nacitarhan S, Tuncer N. Plasma and urine sialic acid in non-insulin dependent diabetes mellitus. Ann Clin Biochem 1995;32(Pt 3):303-6.

8. *Parte de un número*

Poole GH, Mills SM. One hundred consecutive cases of flap lacerations of the leg in ageing patients. N Z Med J 1994;107(986 Pt 1):377-8.

9. *Número sin volumen*

Turan I, Wredmark T, Fellander-Tsai L. Arthroscopic ankle arthrodesis in rheumatoid arthritis. Clin Orthop 1995;(320):110-4.

10. *Sin número ni volumen*

Browell DA, Lennard TW. Immunologic status of the cancer patient and the effects of blood transfusion on antitumor responses. Curr Opin Gen Surg 1993:325-33.

11. *Paginación en números romanos*

Fisher GA, Sikic BI. Drug resistance in clinical oncology and hematology. Introduction. Hematol Oncol Clin North Am 1995 Apr;9(2):xi-xii.

12. *Indicación del tipo de artículo, según corresponda*

Enzensberger W, Fischer PA. Metronome in Parkinson's disease [carta]. Lancet 1996;347: 1337.

Clement I, De Bock R. Hematological complications of hantavirus nephropathy (HVN) [resumen]. Kidney Int 1992;42:1285.

13. *Artículo que contiene una retractación*

Garey CE, Schwarzman AL, Rise ML, Seyfied TN. Ceruloplasmin gene defect associated with epilepsy in EL mice [retractación de Garey CE, Schwarzman AL, Rise ML, Seyfried TN. En: Nat Genet 1994;6:426-31]. Nat Genet 1995;11:104.

14. *Artículo retirado por retractación*

Liou GI, Wang M, Matragoon S. Precocious IRBP gene expression during mouse development [retirado por retractación en Invest Ophthalmol Vis Sci 1994;35:3127] Invest Ophthalmol Vis Sci 1994;35:1083-8.

15. *Artículo sobre el que se ha publicado una fe de erratas*

Hamlin IA, Kahn AM. Herniography in symptomatic patients following inguinal hernia repair [se publica una fe de erratas en West J Med 1995;162:278]. West J Med 1995; 162:28-31.

Libros y otras monografías

(Nota: Con anterioridad, el estilo de Vancouver indicaba, incorrectamente, que entre la editorial y la fecha debía ir una coma en vez de punto y coma, como debe ser.)

16. *Individuos como autores*

Ringsven MK, Bond D. Gerontology and leadership skills for nurses. 2nd ed. Albany (NY): Delmar Publishers; 1996.

17. *Directores (“editores”), compiladores como autores*

Norman IJ, Redfern SI, editors. Mental health care for elderly people. New York: Churchill Livingstone; 1996.

18. *Organización como autor y editorial*

Institute of Medicine (US). Looking at the future of the Medicaid program. Washington: The Institute; 1992.

19. *Capítulo de libro*

(Nota: Con anterioridad, el estilo de Vancouver prescribía el uso de dos puntos en vez de la letra *p* antes de las páginas.)

Phillips SJ, Whisnant JP. Hypertension and stroke. En: Laragh IH, Brenner BM, editors. Hypertension: pathophysiology, diagnosis, and management. 2nd ed. New York: Raven Press; 1995. p. 465-78.

20. *Actas de conferencias*

Kimura I, Shibasaki H, editors. Recent advances in clinical neurophysiology. Proceedings of the 10th International Congress of EMG and Clinical Neurophysiology; 1995 Oct 15-19; Kyoto, Japan. Amsterdam: Elsevier; 1996.

21. *Artículo presentado en una conferencia*

Bengtsson S, Bolheim BG. Enforcement of data protection, privacy and security in medical informatics. En: Lun KC, Degoulet P, Piemme TE, Rienhoff O, editors. MEDINFO 92. Proceedings of the 7th World Congress on Medical Informatics; 1992 Sep 6-10; Geneva, Switzerland. Amsterdam: North-Holland; 1992. p. 1561-5.

22. *Informe científico o técnico*

Publicado por la institución financiadora o patrocinadora:

Smith P, Golladay K. Payment for durable medical equipment billed during skilled nursing facility stays. Final report. Dallas (TX): Dept. of Health and Human Services (US), Office of Evaluation and Inspections; 1994 Oct. Report No.: HHSIGOEI69200860.

Publicado por la institución ejecutora:

Field MI, Tranquada RE, Feasley IC, editors. Health services research: work force and educational issues. Washington: National Academy Press; 1995 Contract No.: AHCPR282942008. Sponsored by the Agency for Health Care Policy and Research.

23. *Tesis doctoral*

Kaplan SI. Post-hospital home health care: the elderly's access and utilization [tesis doctoral]. St. Louis (MO): Washington Univ.; 1995.

24. *Patente*

Larsen CE, Trip R, Johnson CR, inventors; Novoste Corporation, titular. Methods for procedures related to the electrophysiology of the heart. US patent 5,529,067. 1995 Jun 25

Otros trabajos publicados

25. *Artículo de periódico*

Lee G. Hospitalizations tied to ozone pollution: study estimates 50,000 admissions annually. The Washington Post 1996 Jun 21; Sect. A:3 (col. 5).

26. *Material audiovisual*

HIV+/AIDS: the facts and the future [videocassette]. St. Louis (MO): Mosby-Year Book; 1995.

27. *Documentos legales*

Ley pública:

Preventive Health Amendments of 1993, Pub. L. No. 103-183, 107 Stat. 2226 (Dec. 14, 1993).

Proyecto de ley sin sancionar:

Medical Records Confidentiality Act of 1995, S 1360,104th Cong., 1st Sess. (1995).

Código de normas federales:

Informed Consent, 42 C.F.R Sect. 441.257 (1995).

Audiencia:

Increased Drug Abuse: the Impact on the Nation's Emergency Rooms: Hearings Before the Subcomm. on Human Resources and Intergovernmental Relations of the House Comm. on Government Operations, 103rd Cong., 1st Sess. (May 26,1993).

28. *Mapa*

North Carolina. Tuberculosis rates per 100,000 population, 1990 [mapa demográfico]. Raleigh: North Carolina Dept. of Environment, Health, and Natural Resources, Div. of Epidemiology; 1991.

29. *Libro de la Biblia*

The Holy Bible. King James version. Grand Rapids (MI): Zondervan Publishing House; 1995. Ruth 3:1-18.

30. *Diccionarios y obras de consulta semejantes*

Stedman's medical dictionary 26th ed. Baltimore: Williams & Wilkins; 1995. Apraxia; p. 119-20.

31. *Obras clásicas*

The Winter's Tale: act 5, scene 1, lines 13-16. The complete works of William Shakespeare London: Rex; 1973.

Trabajos inéditos

32. *En prensa*

(Nota: La NLM prefiere referirse a estos trabajos como "en preparación" [*forthcoming*] porque no todos se publicarán impresos.)

Leshner AI. Molecular mechanisms of cocaine addiction. N Engl J Med En prensa 1946.

Material en soporte electrónico

33. *Artículo de revista en formato electrónico*

Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis [publicación periódica en línea] 1995 Jan-Mar [citada 1996 Jun 5];1(1):[24 pantallas]. Se consigue en: URL: <http://www.cdc.gov/ncidod/EID/eid.htm>

34. *Monografía en formato electrónico*

CDI, clinical dermatology illustrated [monografía en CD-ROM] Reeves JRT, Maibach H. CMEA Multimedia Group, producers. 2nd ed. Version 2.0. San Diego: CMEA; 1995.

35. *Fichero de computadora*

Hemodynamics III: the ups and downs of hemodynamics [programa de computadora]. Version 2.2. Orlando (FL) Computerized Educational Systems; 1993.

Cuadros

Mecanografíe o imprima cada cuadro a doble espacio y en hoja aparte. No presente los cuadros en forma de impresiones fotográficas. Numérelos consecutivamente siguiendo el orden en que se citan por primera vez en el texto, y asigne un título breve a cada uno. Cada columna llevará un encabezamiento corto o abreviado. Las explicaciones irán como notas al pie y no en el encabezamiento. En las notas al pie se explicarán todas las abreviaturas no usuales empleadas en cada cuadro. Como llamadas para las notas al pie, utilícense los símbolos siguientes en la secuencia que se indica: *, †, ‡, §, ||, ¶, **, ††, ‡‡.

Identifique las medidas estadísticas de variación, tales como la desviación estándar y el error estándar de la media.

No trace líneas horizontales ni verticales en el interior de los cuadros.

Cerciórese de que cada cuadro aparezca citado en el texto.

Si incluye datos publicados o inéditos provenientes de otra fuente, obtenga la autorización necesaria para reproducirlos y conceda el reconocimiento cabal que corresponde.

Incluir un número excesivo de cuadros en relación con la extensión del texto puede ocasionar dificultades al confeccionar las páginas. Examine varios números recientes de la revista a la que planea presentar el artículo y calcule cuántos cuadros pueden incluirse por cada millar de palabras de texto.

Al aceptar un artículo, el director podrá recomendar que los cuadros suplementarios que contienen datos de respaldo importantes, pero que son muy extensos para publicarlos, queden depositados en un servicio de archivo, como el Servicio Nacional de Publicaciones Auxiliares en los Estados Unidos, o que sean proporcionados por los autores a quien lo solicite. En tal caso, se agregará en el texto la nota informativa necesaria. Dichos cuadros se presentarán junto con el artículo para su consideración.

Ilustraciones (figuras)

Envíe los juegos completos de figuras en el número requerido por la revista. Las figuras estarán dibujadas y fotografiadas en forma profesional; no se aceptarán los letreros trazados a mano o con máquina de escribir. En lugar de los dibujos, radiografías y otros materiales de ilustración originales, envíe impresiones fotográficas en blanco y negro, bien contrastadas, en papel satinado y que midan 127 x 173 mm, sin exceder de 203 x 254 mm. Las letras,

números y símbolos serán claros y uniformes en todas las ilustraciones; tendrán, además, un tamaño suficiente para que sigan siendo legibles incluso después de la reducción necesaria para publicarlos. Los títulos y las explicaciones detalladas se incluirán en los pies o epígrafes, no sobre las propias ilustraciones.

Al reverso de cada figura pegue una etiqueta de papel que lleve anotados el número de la figura, el nombre del autor y cuál es la parte superior de la misma. No escriba directamente sobre el dorso de las figuras ni las sujete con broches para papel, pues quedan marcadas. Las figuras no se doblarán ni se montarán sobre cartón.

Las fotomicrografías incluirán en sí mismas un indicador de la escala. Los símbolos, flechas y letras usados en estas deberán contrastar claramente con el fondo.

Si se usan fotografías de personas, éstas no deberán ser identificables; de lo contrario, habrá que anexar un permiso por escrito para poder utilizarlas (véase la sección “Protección del derecho de los pacientes a que se respete su vida privada”).

Las figuras se numerarán en forma consecutiva de acuerdo con su primera mención en el texto. Si la figura ya fue publicada, se reconocerá la fuente original y se presentará la autorización por escrito que el titular de los derechos de autor concede para reproducirla. Este permiso es necesario, independientemente de quién sea el autor o la editorial; la única salvedad son los documentos considerados como de dominio público.

En el caso de las ilustraciones en color, averigüe si la revista necesita negativos, transparencias en positivo o impresiones fotográficas. La inclusión de un diagrama en el que se indique la parte de la fotografía que debe reproducirse puede resultar útil a la redacción. Algunas revistas publican ilustraciones en color únicamente si el autor paga el costo extra.

Pies o epígrafes de las ilustraciones

Los pies o epígrafes de las ilustraciones se mecanografiarán o imprimirán a doble espacio, comenzando en hoja aparte e identificándolos con los números arábigos correspondientes. Cuando se utilicen símbolos, flechas, números o letras para referirse a ciertas partes de las ilustraciones, será preciso identificar y aclarar el significado de cada uno en el pie o epígrafe. En las fotomicrografías habrá que explicar la escala y especificar el método de tinción.

Unidades de medida

Las medidas de longitud, talla, peso y volumen se expresarán en unidades del sistema métrico decimal (metro, kilogramo, litro, etc.) o sus múltiplos y submúltiplos.

Las temperaturas se consignarán en grados Celsius. Los valores de presión arterial se indicarán en milímetros de mercurio.

Todos los valores hemáticos y de química clínica se presentarán en unidades del sistema métrico decimal y de acuerdo con el Sistema Internacional de Unidades (SI). La redacción de la revista podrá solicitar que, antes de publicar el artículo, los autores agreguen unidades alternativas o distintas de las del SI.

Abreviaturas y símbolos

Utilice únicamente abreviaturas corrientes. Evite las abreviaturas en el título y el resumen. Cuando se emplee por primera vez una abreviatura en el texto, irá precedida del término completo, salvo si se trata de una unidad de medida común.

ENVIO DEL MANUSCRITO A LA REVISTA

Envíe por correo el número requerido de copias del manuscrito en un sobre de papel resistente; si es necesario, proteja las copias y las figuras metiéndolas entre dos hojas de cartón para evitar que las fotografías se doblen. Meta las fotografías y transparencias en su propio sobre de papel resistente.

Los manuscritos irán acompañados de una carta de envío firmada por todos los coautores. Es preciso incluir en ella lo siguiente: 1) información acerca de la publicación previa o duplicada, o sobre la presentación de cualquier parte del trabajo a otra revista, según lo expresado líneas arriba; 2) una manifestación de las relaciones financieras o de otro tipo que pudieran dar lugar a un conflicto de intereses (véase más adelante); 3) una declaración de que el manuscrito ha sido leído y aprobado por todos los autores, que se ha cumplido con los requisitos de la autoría expuestos anteriormente en el presente documento y que cada autor está convencido de que el manuscrito representa un trabajo honrado; y 4) el nombre, la dirección y el número telefónico del autor corresponsal, quien se encargará de comunicarse con los demás autores en lo concerniente a las revisiones y a la aprobación final de las pruebas de imprenta. La carta incluirá cualquier información suplementaria que pueda resultar útil para el director, tal como el tipo de artículo que el manuscrito representa para esa revista en particular y si el autor (o los autores) estaría dispuesto a sufragar el costo de reproducir las ilustraciones en color.

El manuscrito se acompañará de copias de los permisos concedidos para reproducir material ya publicado, para usar ilustraciones o revelar información sobre individuos que puedan ser identificados, o para agradecer a ciertas personas su colaboración.

REFERENCIAS

1. United Nations Educational, Scientific, and Cultural Organization. *Bibliography of Publications Designed to Raise the Standard of Scientific Literature*. Paris: UNESCO, 1983
2. Lain Entralgo P. *Historia Universal de la Medicina*. Barcelona: Salvat Editores, 1980
3. Medicina Clínica. *Manual de Estilo*. Barcelona: Ediciones Doyma S.A., 1993
4. Gehlbach SH. *Interpreting the Medical Literature*. New York: Macmillan Publishers, 1982
5. Polit DF, Hungler BP. *Essentials of Nursing Research. Methods and Applications*. Philadelphia: Lippincott Co., 1985
6. Day RA. *Cómo escribir y publicar trabajos científicos*. Publicación Científica 525. Washington, DC: Organización Panamericana de la Salud, 1990
7. CBE Style Manual Committee. *Scientific Style and Format: The CBE Manual for Authors, Editors, and Publishers*. Chicago Sixth edition, Illinois: Council of Biology Editors, 1993
8. Ebel HF, Bliefert C, Russey WE. *The Art of Scientific Writing: From Student Reports to Professional Publications in Chemistry and Related Fields*. Weinheim, Germany: VCH, 1987
9. Jones WP, Keene ML. *Writing Scientific Papers and Reports*. Eighth edition. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1981
10. Silva GA. La autoría múltiple y la autoría injustificada en los artículos científicos. *Boletín de la Oficina Sanitaria Panamericana* 1990; 108(2): 141-152
11. Philbin AI, Presley JW. *Technical Writing: Method, Application and Management*. Albany, New York: Delmar Publishers Inc., 1989
12. Crammins ET. *The Art of Abstracting*. Philadelphia: ISI Press, 1982
13. Teitelbaum H. *How to Write a Thesis: A guide to the Research Paper*. New York: Prentice Hall, 1989
14. Sorgi M, Hawkins C. *Investigación médica. Cómo prepararla y cómo divulgarla*. Barcelona: Ediciones Medici, S.A., 1990
15. Pedhazur EJ, Schmelkin LP. *Measurement, Design and Analysis: An integrated Approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers, 1991
16. *Comité Internacional de Directores de Revistas Médicas*. Requisitos Uniformes para Preparar los Manuscritos enviados a Revistas Médicas. *Boletín de la Oficina Sanitaria Panamericana*. 1994; 116 (2): 146-159

17. Polgar S, Thomas SA. *Introducción a la Investigación en las Ciencias de la Salud*. Madrid: Churchill Livingstone, 1993
18. Barzun J, Graff HF. *The Modern Researcher*. Fifth edition. Fortworth, Texas: Harcourt Brace Publishers, 1992
19. Reynolds L, Simmonds D. *Presentation of Data in Science*. Dordrecht, The Netherlands: 1984
20. World Health Organization. Regional Office for the Western Pacific. *Health Research Methodology. A guide for Training in Research Methods*. Manila: WHO, 1992
21. Howell JB. *Style Manuals of the English-speaking World: A Guide*. Phoenix, Arizona: Oryx Press, 1983
22. Martínez García C., Editor. *El Cáncer en Granada. Incidencia y Mortalidad. 1988-90* Granada: Escuela Andaluza de Salud Pública, 1994
23. Martín Andrés A, Luna del Castillo J de D. *Bioestadística para las Ciencias de la Salud*. 4ª Edición. Madrid: Ediciones Norma, 1994
24. Armitage P, Berry G. *Statistical Methods in Medical Research*. Oxford: Blackwell Scientific Publications, 1987
25. Tufte ER. *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press, 1983
26. Turabian KL. *A Manual for Writers of Term Papers, Thesis, and Dissertations*. Fifth edition. Chicago: The University of Chicago Press, 1987
27. *The Chicago Manual of Style*. Fourteenth edition. Chicago: The University of Chicago Press, 1993
28. Rothman KJ. *Epidemiología Moderna*. Madrid: Ediciones Díaz de Santos, S.A., 1987
29. Sackett DL, Haynes RB, Tugwell P. *Epidemiología Clínica*. Madrid: Ediciones Díaz de Santos, S.A., 1989
30. Ramón y Cajal S. *Reglas y Consejos sobre investigación científica*. Madrid: Espasa Calpe, 1991
31. Medewar PB. *Consejos a un joven científico*. México: Fondo de Cultura Económica, 1982
32. Real Academia Española. *Diccionario de la Lengua Española*. Madrid: Real Academia Española, 1992
33. Gibaldi J, Ahtert WS. *MLA Handbook for Writers of Research Papers*. Third edition. New York: The Modern Language Association of America, 1988
34. Cordera A, Bobenrieth M. *Administración de Sistemas de Salud*. Tomo II. Mexico: Cordera, 1983
35. Mulrow CD. El Artículo de Revisión en la literatura Médica Actual. *Boletín de la Oficina Sanitaria Panamericana*. 1993; 114(5): 437-459
36. Delahaye F, Landrison G, Ecochord R, Colin C. Meta-analysis *Health Policy*. 1991; (19): 185-196
37. Jones DR. Meta-analysis of Observational Epidemiological Studies: a Review. *Journal of the Royal Society of Medicine*. 1992. 85: 165-168

38. Martín Vivaldi G. *Curso de Redacción. Teoría y Práctica de la Composición y del Estilo*. Madrid: Paraninfo, 1986
39. Miranda Podera L. *Gramática Española*. 12 Edición. Madrid: Editorial Hernando, 1982
40. Seco M. *Gramática esencial del Español. Introducción al estudio de la lengua*. Madrid: Editorial Espasa Calpe, S.A., 1994
41. Alarcos Llorach E. *Gramática de la lengua española*. Madrid: Editorial Espasa Calpe S.A., 1994
42. Baena G. *Redacción Práctica. El Estilo Personal de Redactar*. Mexico: Editores Mexicanos Unidos, 1993
43. Lapesa R. *Introducción a los estudios literarios*. 16 Edición . Madrid: Cátedra, 1986
44. Ferrater Mora J. *Diccionario de Filosofía Abreviado*. Barcelona: Edhasa, 1993
45. Martínez de Sousa J. *Diccionario de redacción y estilo*. Madrid: Ediciones Pirámide SA., 1993
46. Rojas Tapia A. *Manual de Redacción Lógica*. Mexico: Ediciones Quinto Sol, 1988
47. Miranda Podadera L. *Curso de Redacción*. Madrid: Editorial Hernando, 1985
48. Martínez de Sousa J. *Dudas y errores de lenguaje*. Cuarta edición. Madrid: Paraninfo, 1987
49. Delton, J. *The 29 most common writing mistakes and how to avoid them*. Cincinnati, Ohio: Writers Digest Books, 1985
50. Martínez Amador E. *Diccionario Gramatical y de Dudas del Idioma*. Barcelona: Editorial Ramón Sopena, S.A., 1970
51. Winkler J.T. El Síndrome de la celebridad intelectual. *The Lancet* 1987; 11(1):65
52. Andersen B. *Methodological Errors in Medical Research*. Oxford: Blackwell Scientific Publications, 1990
53. Downs F., Newman M. *In A Source Book of Nursing Research*. Third edition. Philadelphia: F.A. Davis Co., 1983
54. Polit DF, Hungler B.P.. *Investigación Científica en Ciencias de la Salud*. Tercera edición. Mexico: Interamericana/Mc Graw-Hill, 1991
55. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991
56. Sproul N. *Handbook of Research Methods*. London: The Scarecrow Press Inc., 1988
57. Chubin DE, Hackett EJ. *Peerless Science. Peer Review and U.S. Science Policy*. Albany, New York: State University of New York Press, 1990
58. Council of Biology Editors. *Peer Review in Scientific Publishing*. Chicago, Illinois: Council of Biology Editors, Inc., 1991
59. Lock S. *A Difficult Balance. Editorial Peer Review in Medicine*. Philadelphia: ISI Press, 1985
60. Silva G, Campillo Artero C. Cómo se deben evaluar los artículos científicos propuestos para publicación. *Medicina Clínica* (Barcelona) 1991; 744-48

ÍNDICE DE MATERIAS

A

Abreviación, 333
Abreviatura, 326, 328, 329, 356, 383
Accidentes gramaticales, 438
Actas
 anatomopatológicas, 306
 de conferencias clínicas, 306
Actas Urológicas, 308
Agradecimientos, 324, 325, 331, 413-414
Alcance del estudio, 339
Aleatoria, 36, 43, 178, 262
Alerta bibliográfica, 329
Álgebra de Boole, 298
Ambiguo (título), 326
Amplitud de intervalo, 136
Ampuloso, 326
Análisis
 crítico de una investigación, 467
 crudo, 94
 estratificado, 48
 multivariante, 234-235
 sensibilidad, 268-269
 supervivencia, 123, 219-228, 247-253
 textual, 288
 varianza (ANOVA), 212
Annals of Internal Medicine, 308
Antecedentes del problema, 337
Anuarios, 307
Año de publicación, 327
Apareamiento, 47, 104
Apéndice(-s), 325, 424-426
Arbitraje por pares, 306
Árbol de decisión, 260-268

Archivos Españoles de Urología, 308
Artículo científico original, 309, 310, 325,
 analítico, 325
 aplicado, 325
 descriptivo, 325
 elementos, 323-324
 experimentales, 325
 no experimentales, 325
 partes del cuerpo, 324, 325
 partes finales, 324, 325
 partes preliminares, 324, 325
 teóricos, 325
Artículo de revisión, 309, 427-430
Artículo de revista, 421, 423
ASCII, 291
Asignación aleatoria, 36, 111, 119
Asociación, 253-258, 319
Atención Primaria, 308
Autor, 324, 330-333
Autoría, 332
 incompleta, 331
 injustificada, 331
 irresponsable, 331
 múltiple, 331
Azar, 161

B

Backward stepwise, 242
Base de estudio, 59, 80, 95, 103, 109
Bases de datos, 298
Bases para presentación tabular, 352
Bibliografía, 324
Biblioteca Nacional de Medicina, 391

- BMDP, 300
 Boletines, 307
Boletín de la Oficina Sanitaria Panamericana, 307, 462
Bulletin of WHO, 307
British Medical Journal, 132, 308
- C**
- CA Bibliofile, 327
 Caja, 154, 158
 Cambio relativo, 395
 Campo, 280, 298
 codificado, 287
 definición de, 285
 derivado, 289
 especial, 286
 lógico, 286
 longitud (anchura), 287
 nombre del, 285
 tipo de, 286
 Capítulo de libro, 421, 423
 Caras de Chernoff, 146
 Carta al editor, 310
 Caso clínico, 309
 Caso prevalente, 83, 96, 99
 Causalidad, 253-258
 Centrado de variables, 252
Chemical Abstract Services, 327
 Ciego, 123
 Claridad, 291, 359
 Coeficiente, 211
 correlación, 211
 determinación, 211
 Pearson, 209
 regresión lineal, 206, 209, 238-241
 regresión logística, 216-218
 Spearman, 212, 213
 variación, 154
 Coherencia, 258
 Cohorte, 39, 78, 102
 abierta (dinámica), 54, 59, 81, 103
 cerrada (fija), 54, 79, 103
 Coma, 407
 Comillas, 407
 Comité Internacional de Editores de Revistas Médicas (CIERM), 332, 343, 391, 489
 Comparabilidad de dos grupos, 192
 Comparación intermedia, 124
 Comparaciones múltiples, 201
 Compendio, 306
 Comunicación breve, 309
 Comunicación científica, 305
 Concisión, 329, 436
 Conclusión, 323, 324, 325, 338, 346, 410, 412-413
 Confiabilidad, 346
 Confusión, 229-231
 Confuso (título), 327
 Conocimiento
 características, 314-317
 científico, 314-317
 concepto, 314
 cotidiano, 314
 teoría del, 314
 Consentimiento informado, 122
 Consistencia, 257
 Control, 341
 Control de la asignación, 35, 111
 Correlación, 67, 208-213
 Credencialismo, 331
 Criterios
 autoría, 330
 exclusión, 342
 inclusión, 342
 Cualidades del buen estilo, 435-437
 Cuartiles, 132, 150
 Cuasiexperimentos, 330, 384
 Cuerpo de la tabla, 315, 354
 Cuerpo del artículo, 324, 325
 Curvas ROC, 154-156, 172-174
- D**
- Datos, 279, 280
 análisis y tratamiento, 297
 diccionario, 285, 290
 fichero, 283
 entrada, 279
 estructura de, 281
 formato, 290
 fuente, 281
 información, 279-293
 preparación, 279-293

protocolo de recogida, 277, 279
protocolo de tratamiento, 277
soporte, 281

DBASE, 286, 291

DBASE III, 292

DBASE IV, 292

Definición
términos, 337
operacional de términos, 320, 339, 342
variables, 337

Denominadores, 31, 49

Densidad de incidencia (tasa), 55, 59, 60, 80

Descriptores, 327

Desviación media, 152

Desviación típica, 153

Diagonal, 355, 431

Diagrama
barras, 138
estrellas, 145
sectores, 140

Diccionario, 306

Diéresis, 407

Directorio, 306

Discusión, 323, 324, 383

Diseño, 320, 341
casos, 320
control, 342
cuasiexperimental, 320
ecológicos, 16, 20, 34, 38, 47, 65
epidemiológicos, 320
experimental, 320
gráficos, 371
gráficos en barra, 374
gráficos lineales, aritmética, 387
gráficos lineales, semilogarítmica, 396
gráficos en sectores, 399
gráficos de dispersión, 402
histogramas, 382
longitudinal, 320
métodos, 341
nivel I descriptivo, 320
nivel II analítico, 320
nivel III inferencial, 320
no experimental (observacional), 320
polígonos de frecuencia, 385
tablas, 352
tipos, 35
transversal (seccional), 320
validez, 342

Dispersión (gráfico), 401

Documentales, 342

Dos puntos, 431

E

EDIT, 291, 297

Editores de texto, 295-297

Editorial, 310

Efectividad, 247, 265

Elegibilidad de sujetos, 343

Empirismo, 317

Encabezamientos de columnas, 354

Enciclopedia, 306

Enmascaramiento, 123

Ensayo clínico, 36, 116, 309

Epidemiología, 10, 35, 43

EPIINFO, 291, 300

EPILOG, 300

Epílogo, 334, 338

Error, 43
aleatorio, 43, 178, 255
alpha, 172, 190
beta, 172, 190
estándar, 162, 180
estándar de la media, 180
estándar de una producción, 181
sistemático (sesgo), 44, 83, 104, 121

Escala aritmética, 372

Escala logarítmica, 376

Escritura, 303

Especificidad, 162, 168-170, 248

Estándar ideal, 162

Estilística, 432

Estilo, 328, 332, 432
científico, 433
científico y estilo literario, 433
de nombre-año (o de Harvard), 422
de orden de mención (o numérico), 420
en la comunicación científica, 431
literario, 433

Estimación, 179

Estimador, 86, 179

Estratificación, 231

Estrato, 233

- Estudio, 322
 alcance, 322
 aleatorio controlado, 36, 108
 aportes, 322
 caso-control, 40, 75, 95
 conclusión, 322
 correlacional, 325
 cuasiexperimental, 37, 74, 112
 de seguimiento, 39, 74, 78
 descriptivo, 34, 37, 63
 ecológico, 16, 20, 34, 38,47, 65
 experimental, 18, 36, 93, 111
 observacional (no experimental), 19, 37
 piloto, 320
 preclínico, 99, 117
 transversal, 16, 21, 34, 39, 51, 69, 95, 113
- Etapas del proceso de investigación, 311, 318-323, 324
 conceptual (teórica), 319, 324
 diseminación (publicación), 322, 324
 empírica (práctica), 320, 324
 interpretativa (significado), 321, 324
 planeación (diseño), 320, 324
- Ética, 307, 345
- Evaluación
 artículos médicos generales, 485
 artículos originales, 463
 comunicados de investigación, 469
 crítica de la investigación publicada, 477
 en la investigación, 481
 informes de ensayos clínicos, 487
 lista de comprobación, 479
- Excell, 297
- Excerpta Médica*, 334
- Experimentación, 36
- Experimentos verdaderos, 384
- Extenso, 326
- Expurgo de fuentes primarias, 459
- F**
- Falacia ecológica, 68
- Falsos negativos, 163-165, 168-170
- Falsos positivos, 163-165, 168-170
- Fase
 conceptual, 336,
 empírica, 345
 interpretativa, 408
- Fichero, 280
- Figuras, 348
- Formato de datos, 269
- Fórmula, 333
- Formulación de hipótesis, 338
- Forward stepwise, 223, 241
- Fotografía, 50, 320
- Fracción etiológica, 93, 109
- Frecuencia, 116, 134
 absoluta, 116, 134
 acumulativa absoluta, 136
 acumulativa relativa, 136
 relativa, 116, 134
- Fuentes de información, 304, 317, 342
 primaria, 458, 459
 secundaria, 458
- Fuerza de asociación, 67, 85, 87, 105, 209
- Función de riesgo, 229, 247
 de riesgo base, 230, 248
 de supervivencia, 203, 221
- G**
- Gaceta Sanitaria*, 308
- Gestor de base de datos, 295, 298-299
- Glim, programa estadístico, 300
- Gráficos, 136-147, 370-405
 barra, 138-139, 374-381
 barras agrupadas, 374
 barras componente 100%, 374
 barra horizontal, 375
 barra simple, 374
 barra subdividida (segmentada), 374
 barra vertical, 375
 cambio relativo, 395
 caras de Chernoff, 144, 146
 columna, 374
 curvas epidémicas, 369, 384
 datos bidimensionales, 142-144
 datos multidimensionales, 144-147
 datos unidimensionales, 138-142
 desviación, 374
 diagrama de dispersión, 401
 diagrama de estrellas, 126, 127
 diseño y construcción, 371
 dispersión, 401-405

histogramas, 161, 381-384
 lineal de coordenadas rectilíneas, 386
 lineal de escala aritmética, 386-394
 lineal de escala semilogarítmica, 395-398
 nube de puntos, 143
 objetivos, 137, 370
 pareados, 374
 pictográficos, 374-375
 polígono de frecuencias, 139, 385-386
 polígono de frecuencias acumulativas
 absolutas, 140
 principios básicos, 371
 proporción, 395
 sectores (circular, en tarta) 122, 399-401
 técnica de cero suprimido, 387
 tronco y hojas, 141-142

H

Hallazgos principales, 345
 Hipótesis, 34, 319, 339, 410
 Hipótesis nula, 187
 Histograma, 139, 161, 363, 452
 Hojas de cálculo, 295, 297-298

I

IBID, 407
 IBIDEM, 407
 Idealismo, 317
 Impacto potencial, 83, 319
 IMRYD, 323
 Incidencia, 50, 52, 68, 96, 102, 319
 Incidencia acumulada (riesgo), 53, 79, 107, 109
Index Medicus, 307, 319, 334, 415, 420
 Indicador, 61, 64, 112
 Índice, 44, 62
 Inferencia estadística, 178
 Información, 279-280, 320
 análisis, 321
 confiabilidad, 320
 fuentes, 320
 presentación, 321
 primaria, 306, 428

procesamiento, 321
 secundaria, 428
 validez, 320

Informática, 277

Informe

de caso, 38, 64
 de congreso, 309
 de jornadas, 309
 técnico, 307

Instrumentos, 342

Interacción, 231-233

Interrogación, 431

Intervalo

de clases, 354-355
 de confianza, 44, 182-185, 343
 de confianza para diferencia entre
 proporciones, 173-174, 191-192

Intervención, 35, 74, 109, 111, 342

Introducción, 323, 324, 336-340

Investigación, 311-313

análisis de contenido, 313
 concepto, 311
 comprobación para evaluar, 479-480
 comunicación, 305
 definiciones, 311-312
 diseños, 31-41
 elementos para análisis crítico, 467-468
 etapas, 318-323
 evaluación de comunicados, 469-475
 lista de evaluación, 481-484
 métodos/metodología, 21, 317-318
 principios, 317
 proceso, 317-324
 soporte informático, 277
 tipo causal, 328
 tipo descriptivo, 328

J

Jerga, 328, 329

Jerigonza, 329

K

Kaplan-Meier (método), 221-224

L

Lectura crítica de artículos científicos, 303, 445-451

Ley de azar, 161

Libros, 306, 421, 422

- de texto, 306
- de referencia, 306
- reseñas, 310

Libros y monografías, 397

Limitaciones del estudio, 339

Lista de autores, 327

Listas de revisión, 446

Literatura biomédica, 305

- primaria, 305, 427
- secundaria, 305, 427

LOC.CIT, 408

LOCO CITATO 408

LOGRESS, 300

Lotus, 291

Lotus 1-2-3, 297

M

Magnitud de asociación, 85, 105, 123, 255-256

Manipulación de una variable, 36

Manual, 306

Mapas, 364, 365

Marca de clase, 118, 136

Marco conceptual, 319

Marco teórico, 319, 337, 339, 408

Materiales y métodos, 324, 341-345

Materiales y técnicas, 341

Matriz, 281-283

- dentada, 283
- de tablas, 361, 362
- en dos, 283
- fichero, 283
- rectangular, 281, 282

Máxima verosimilitud, 215

Media, 147

Media ponderada, 149

Mediana, 150

Medidas de frecuencia, 49

Medicina Clínica, 308

MEDLARS, 327

MEDLINE, 307

Método, 317, 341

Método estadístico, 342

Metodología, 317-318

Microsoft Word, 296

MICROSTAT, 300

Mínimos cuadrados, 205

Modelo, 247

- Cox, 247-253
- regresión, 202
- regresión lineal, 202
- riesgos proporcionales, 253

Monografía, 306

MS-DOS, 297

Muestra, 177, 320

- independiente, 320
- relacionada, 320

Muestreo, 178, 320

- no probabilístico, 178, 320
- probabilístico, 178, 320

N

Nivel de confianza, 182

Nivel de error, 183

Nodos, 261

- de decisión, 261
- probabilísticos o aleatorios, 262
- terminales, 262

Nombre de la revista, 327

Nombre y año (estilo Harvard), 417

Notas

- contenido, 405
- pie de página, 405
- pie de tabla, 358
- referencia, 405

Noticia, 309

Novedad, 346

Nube de puntos, 143

Numeración de gráficos, 371

Numeración de tablas, 353

Numérico-alfabético, 417

O

Objetivo(s), 319, 338, 339, 409

Observación, 37

rara, 151

extremas (outliers), 208

influyentes, 208

Odds, 89, 105

Odds Ratio, 73, 89, 102, 105

Oncología, 308

OP. CIT. 407, 408

OPERE CITATO 407

OR, 73, 89, 102, 105

Orden de mención, 417

Ordenador, 279

Osiris, 300

P

Pacientes y métodos, 340

Palabra,

clave, 326, 327

correcta, 327

Paquete estadístico, 296, 299

Paréntesis, 355

Párrafo, 356

Parte bibliográfica del artículo, 327

Partes del artículo científico original, 325-426

cuerpo, 325

finales, 325

preliminares, 325

Partes del artículo de revisión, 427

Pautas

de evaluación, 447

para la calidad de revisión, 428

para la evaluación de artículos originales, 463

Película, 50

Percentiles, 150

Permuterm Subject Index, 327

Peroración de un discurso, 412

Personas-tiempo, 56, 81

Pictográficos, 374

Planeación, 320

Plausibilidad biológica, 257

Población, 177, 320

accesible, 409

blanco, 409

diana, 409

en estudio, 54, 60, 95, 102

Poder, 342

Polígono de frecuencias, 139

Polígono de frecuencias acumulativas

absolutas, 140

Ponencia a congresos, 306

Positivismo, 317

Potencia de un test, 190

Práctica, 325

Precisión, 342, 412

Predictoras categóricas, 237

Pregunta

abierta, 288

investigación, 341, 347, 410

multirespuesta, 288

Premio Nobel, 314, 315

Prevalencia, 49, 51, 69, 93, 319

Primera revisión bibliográfica, 337

Principios, 317

Probabilidad, 52, 161

condicionada, 161, 162

postest, 165

pretest, 165

promedio individual, 52

umbral, 269

Problema, 299, 337

de investigación, 319, 337, 339

general, 319, 338, 339

Procedimientos, 342

Procesador de textos, 295-297

Proceso de aleatorización, 343

Proceso de datos, 279

Proporción, 51

atribuible, 92, 109

atribuible poblacional, 93, 109

Prospectivo, 78

Protocolo de tratamiento de datos, 277

Protocolo de recogida de datos, 279

Pruebas de comprobación de hipótesis, 343

Publicaciones científicas, 305

en revistas, 324

tipos, 322

Publicaciones periódicas, 305

Publicaciones de tipo primario, 337
 Punto, 431
 Punto y coma, 431
 Puntos suspensivos, 431

Q

Qedit, 296
 Quattro-Pro, 297

R

Racionalismo, 317
 Rango, 287
 Razón
 prevalencia, 72
 riesgo, 86
 tasas, 87
 Realismo, 317
 Recorrido intercuartílico, 154
 Redacción de títulos, 453-456
 Redacción, gramática y estilo, 431
 Referencias, 325, 415-424
 Registro, 280, 298
 diseño, 280
 estructura, 281, 284
 unidad, 281
 Regla de Bonferroni, 201
 Regresión, 197
 a la media, 114
 lineal múltiple, 236-242
 lineal simple, 197, 202-208
 logística, 197
 logística múltiple, 242-246
 Relación dosis-respuesta, 256
 Relación temporal, 255
 Requisitos de una buena ilustración, 348
 Requisitos de una buena tabla, 309, 347
 Requisitos uniformes para preparar manuscritos
 489
 Reseña de libros, 310
 Residuales, 208
 Respuesta de investigación, 412
 Restricción, 47

Resultados, 321-322, 323, 324
 Resumen(es), 324, 325, 333-336
 artículo, 333
 estructurado, 334
 ponencia, 333
 tradicional, 334
 Retrospectivo, 78
 Review, 291
 Revisión, 305
 literatura, 467, 469, 481
 literarias publicadas, 427
 por árbitros, 306, 340
 Revisión bibliográfica, 319, 337, 457-462, 463,
 467, 477, 479
 búsqueda bibliográfica, 458
 evaluación de las fuentes, 460
 lectura crítica, 459
 momentos, 457
 objetivos, 457
 Revistas, 305-306
 biomédicas españolas, 308
 científicas, 307
 enfermedades, 305
 especialidades, 305
 investigación, 305
 literatura primaria, 305
 literatura secundaria, 305
 modalidades de atención, 305
 profesiones, 305
 secciones comunes, 309
Revista de Calidad Asistencial, 308
Revista Clínica Española, 308
Revista de Sanidad e Higiene Pública, 308
Revista Panamericana de Salud Pública 489
 Riesgo, 52, 319
 relativo, 86, 90, 107, 319
 atribuible, 91, 109, 319
 atribuible poblacional, 93, 109
 Rigor metodológico, 277

S

Science Citation Index, 327
 Segunda revisión bibliográfica, 409, 410
 Selección de variables, 241
 Selección del tipo de gráfico, 348
 Semilogarítmico (gráfico), 395-398

- Sencillez y Naturalidad, 437
- Sensibilidad, 162, 168-170
- Serie
- casos, 38, 65
 - continuada, 306
 - temporal, 67, 113
- Sesgo, 342, 409
- Berkson, 103
 - clasificación, 44, 84, 104, 123, 342
 - confusión, 45, 104, 342, 409
 - diferencial, 44, 85, 104
 - información, 409
 - no diferencial, 44, 85, 104
 - selección, 44, 71, 83,, 103, 122, 342, 409
- Shareware, 297
- Sigla, 326, 329
- Sigma, programa estadístico, 300
- Significación, 346
- Significación práctica, 193
- Significado, 408
- Signos de admiración, 431
- Sintaxis correcta, 327
- Sistema Vancouver, 415
- Sobreexplicación, 329
- Software, 295
- Software estadístico, 300
- Soporte informático, 277
- SPAD, 300
- SPAD-T, 300
- S-PLUS, 300
- SPSS/PC, 300
- Stat-Graphics, 300
- Stepwise, 241-242
- Subtítulo, 328
- Sucesos
- incompatibles, 161
 - independientes, 161
- Sujetos, 342
- Sujetos y métodos, 341
- Suplemento con hojas sueltas, 306
- T**
- t de Student, 197, 198-202
- Tablas, 133-136, 347, 348, 350-369
- contingencia (cruzada, dos por dos) 186, 362-364
 - diseño y construcción, 352
 - dos variables, 361-362
 - intervalos de clases, 354-357
 - matriz de (formato o plantilla), 367-369
 - objetivos, 347
 - otras medidas estadísticas, 366
 - principios básicos, 352
 - requisitos, 348
 - tres o más variables, 364-365
 - una variable, 360-361
- Tamaño
- gráficos, 364
 - muestra, 44, 177
 - muestra para diferencia entre proporciones, 193-196
 - tabla, 357-358
- Tasa de riesgo, 247
- Tasa instantánea, 60
- Teorema de Bayes, 165
- Teoría, 312, 325
- Teoría del conocimiento, 314
- Test, 197
- diagnósticos, 161-176
 - Ducan, 202
 - hipótesis, 44, 187-191
 - Log-Rank, 224-228
 - Mann-Whitney, 201
 - Mantel-Haenszen, 224-228
 - Newmann-Keuls, 202
 - no paramétrico, 201
 - Tukey, 202
 - Wilcoxon, 201
 - significativo, 189
- Texto, 347
- The Journal of the American Medical Association, 308
- The Lancet, 307
- The New England Journal of Medicine, 132, 193, 307
- Tiempo de supervivencia, 220
- Tipos de estudio, 31, 35
- Titulillo, 328
- Título, 324, 325, 326-330
- ambiguo, 326
 - confuso, 326
 - con subtítulos, 329
 - efectista, 329, 353

gráfico, 364
llamativo, 353
página, 328
partido, 329
revista, 328
tabla, 353

Tono efectista, 326

Transformación logit, 214, 243

U

Unidad de observación, 320

Unidad de registro, 281, 282

Utilidad, 260, 268, 269

V

Validez, 322

externa, 30, 48, 322, 341, 342, 408, 409

interna, 44, 236-238, 254-256, 341, 342, 408

Valor

adyacente, 154, 155

faltante, 287

P, 189, 343

predictivo negativo, 167

predictivo positivo, 167

registro, 280

Variable, 131, 319, 339

alfanumérica, 286

categoría, 337

confusión, 45, 94, 114, 116, 231, 338

continua, 319

control, 319, 338

cualitativa, 319

cualitativa dicotómica, 133

cualitativa ordinal, 133

cualitativa policotómica, 133

cuantitativa, 319

cuantitativa continua, 133

cuantitativa discreta, 133

dependiente, 197, 319, 338

discreta, 319

dummy, 337

independiente, 197, 206, 319, 338

interviniente, 319, 338

linealmente independientes, 206

nominal, 319

numérica, 286

ordinal, 319

predictora, 197

respuesta, 197

Varianza, 153

Verdaderos negativos, 163-165

Verdaderos positivos, 163-165

Vicios de mal estilo, 438-444

W

Word Perfect, 296

Word Star, 296

ÍNDICE DE AUTORES

- Albalat, A., 436
Alhbom, A., 31
Allen, W., 13
Altman, D.G., 186, 485, 487
Aristóteles, 317
Azorín [José Martínez Ruiz], 316
- Bacon, Francis, 315
Baena, Guillermina, 312
Baroja, Pío, 436
Bartholin, Thomas, 17
Bayes, Thomas, 165
Blegny, Nicolas de, 17
Blouin, Eglá, 433
Bobenrieth Astete, Manuel A., 305, 311, 325, 403, 421, 429, 463
Bradford Hill, A., 82, 253
Breslow, N.E., 96
Brink, Pamela J., 312
Burgos Rodríguez, Rafael, 11, 36, 159, 259
Burns, Nancy, 312
- Campbel, M.J., 196
Cervantes, Miguel de, 440
Chejov, Anton, 436
Chicharro Molero, José A., 159, 259
Christensen, E., 249
Clemanceau, George, 440
Comte, Augusto, 316, 317
Curie, Marie, 314, 316
- Day, N.E., 96
Descartes, René, 11, 317
Doll, R., 82
Downs, Florence, 467
- Einstein, Albert, 317
Eliot, George [Mary Ann Evans], 420
Elwood, J.M., 254
- Fernández Romacho, Irene, 433
- Gadner, M.J., 186
García Puche, José L., 14
Gracia, Diego, 12
Grove, Susan K., 312
- Hott, L.F., 312
Hungler, Bernadette, 312, 469
- Kant, Emmanuel, 314
Kerlinger, Fred N., 314
- Laín Entralgo, Pedro, 314
Leclerc, George L., 433
Leedy, Paul D., 312
Leibniz, Wilhelm von, 317
- Machado, Antonio, 439
Machín, D., 196, 227
Massey, Veta H., 312
Mosby, M., 312
Moses, L.E., 131
- Newman, Margaret, 467
Nietzsche, Friedrich, 437
Nightingale, Florence, 137
Norel, S., 32
Notter, Lucile E., 312
- Ocaña Riola, Ricardo, 131, 229
Ortega y Gasset, José, 435, 440
- Parménides, 317
Pasteur, Louis, 18
Pavlov, Ivan, 314
Perea-Milla López, Emilio, 31, 43, 49, 63, 77, 95, 111, 229
Playfair, J., 137

Polgar, Thomas, 477, 479
Polit, Denise F., 312, 469
Popper, Karl, 13
Pringle, M., 236

Ramón y Cajal, Santiago, 315
Renan, Ernest, 432
Ross, Robert, 312
Rothman, Kenneth J., 54
Russell, Bertrand, 315, 316

Sánchez-Cantalejo Ramírez, Emilio, 131, 177, 197,
229
Sallo, Denis de, 17
Shakespeare, William, 436

Taber's Encyclopedía, 312
Tannery, Jules, 316
Tukey, J.W., 141

Unamuno, Miguel de, 414

Voltaire [François Marie Arouet], 433

Ware, J. H., 190
Webster Dictionary, 312
Wilde, Oscar, 440
Winkler, J.T., 444
Wood, Marilyn J., 312